

Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós Graduação em Economia

Ensaio sobre Modelos de Previsão Econômica

Lucas Lúcio Godeiro

João Pessoa - PB
2018

Lucas Lúcio Godeiro

Ensaio sobre Modelos de Previsão Econômica

Tese apresentada ao Programa de Pós
Graduação em Economia da Universidade
Federal da Paraíba - UFPB, em cumprimen-
to às exigências do curso de Doutorado
em Economia.

Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós Graduação em Economia

Orientador: Dr. Luiz Renato Régis de Oliveira Lima

João Pessoa - PB
2018

Catálogo na publicação
Seção de Catalogação e Classificação

G581e Godeiro, Lucas Lúcio.

Ensaio sobre Modelos de Previsão Econômica / Lucas
Lúcio Godeiro. – João Pessoa, 2018.
116 f.

Orientação: Luiz Renato Régis de Oliveira Lima.
Tese (Doutorado) – UFPB/CCSA/PPGE.

1. Incerteza Macroeconômica. 2. Regressão de Texto. 3.
Aprendizado de Máquina. 4. Previsão Macroeconômica. 5.
Previsão Financeira. I. Lima, Luiz Renato Régis de
Oliveira. II. Título.

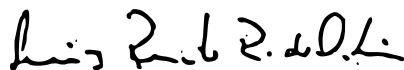
UFPB/BC

Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós Graduação em Economia

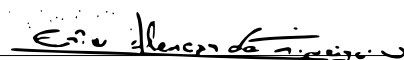
Comunicamos à coordenação do curso de Pós Graduação em Economia que a tese de Doutorado do Aluno Lucas Lúcio Godeiro, matrícula 2015103607 foi submetida à apreciação da comissão examinadora elencada abaixo; no dia ____/____/____ às ____.

A tese foi _____ pela comissão examinadora.

Reformulações sugeridas: Sim () Não ()



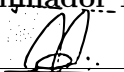
Dr. Luiz Renato Régis de Oliveira Lima
Orientador



Dr. Erik Alencar de Figueiredo
Examinador Interno



Dr. Cássio da Nóbrega Besarria
Examinador Interno



Dr. Márcio André Veras Machado
Examinador Externo ao Programa



Dra. Raquel Menezes Bezerra Sampaio
Examinadora Externa à instituição

Agradecimentos

Agradeço a Deus por me dar forças e chegar ao final de mais uma etapa da minha vida.

Agradeço à minha esposa Beônia, e ao meu filho Jailson por sempre estarem ao meu lado dando apoio e carinho. Também agradeço à minha mãe Fátima e ao meu irmão Vinícius. Agradeço também à Ednalba(Babá), que é considerada da família, por estar a tantos anos cuidando da gente. Em nome do meu sogro Chiquinho e da minha sogra Anelita agradeço a toda a família da minha esposa Beônia, por todo apoio durante essa jornada.

Agradeço ao meu orientador Luiz Renato Lima por todos os ensinamentos transmitidos e suporte para a construção desse trabalho, bem como pela oportunidade de cumprir 1 ano do meu curso como Visiting Scholar na Universidade do Tennessee. Agradeço também à sua esposa Ingrid Lima pelo apoio e amizade prestados a mim e a minha família durante nossa estadia em Knoxville/TN.

A University of Tennessee por ter me recebido durante 1 ano como visiting scholar, em especial ao meu co orientador Mohammed Mohsin. Também agradeço à CAPES por financiar a minha experiência como visiting scholar.

Aos professores do Programa de Pós Graduação em Economia da UFPB, em especial aos professores Cássio Nóbrega e Erik Figueiredo que se dispuseram a fazer parte da comissão examinadora. Também agradeço aos secretários Ricardo e Risomar por todo apoio durante o curso.

Agradeço ao professor Márcio Machado e a professora Raquel Menezes por participarem da comissão examinadora.

Resumo

Esta tese compreende 3 capítulos sobre o tema Modelos de Previsão Econômica. Abaixo seguem os resumos individuais de cada capítulo.

Capítulo 1: **Medindo Incerteza Macroeconômica para o Brasil**

Este capítulo propõe uma nova medida de incerteza macroeconômica para o Brasil. O índice foi construído com base na metodologia de [Jurado et al. \(2015\)](#), utilizada na construção do índice para a economia americana. Mostramos que o aumento no nível de incerteza precede as duas últimas recessões no Brasil e que a mudança na política macroeconômica ocorrida em 2010-2011, elevou de forma substancial o nível de incerteza na economia brasileira. Nosso trabalho sugere que a medida proposta de incerteza possui potencial para analisar políticas macroeconômicas e prever recessões.

Capítulo 2: **Identificando o Poder Preditivo das Minutas do FED**

Este capítulo propõe um novo método para extrair as informações mais preditivas das minutas do FED. Em vez de considerar um dicionário (conjunto de palavras) com um conteúdo fixo, construímos um dicionário cujo conteúdo pode mudar com o tempo. Especificamente, utilizamos o aprendizado de máquina(*Machine Learning*) para identificar as palavras mais preditivas (o conteúdo mais preditivo) de uma determinada minuta e usá-las para derivar novos preditores. Mostramos que os novos preditores melhoram a acurácia da previsão do crescimento do Produto por uma margem estatisticamente significativa, sugerindo que a combinação de aprendizado de máquina e regressão de texto(*Text Regression*) pode ser interpretada como um dispositivo poderoso para a previsão macroeconômica fora da amostra.

Capítulo 3: **Previsão do Prêmio de Risco:Identificando o Poder Preditivo das Notícias Financeiras**

Este capítulo propõe um novo método para extrair as informações mais preditivas das notícias financeiras do Wall Street Journal e do The New York times. Em vez de considerar um dicionário (conjunto de palavras) com um conteúdo fixo, construímos um dicionário cujo conteúdo pode mudar com o tempo. Especificamente, utilizamos o aprendizado de máquina(*Machine Learning*) para identificar as raízes de palavras/termos mais preditivos (o conteúdo mais preditivo) de uma determinada notícia e usá-las para derivar novos preditores. Mostramos que os novos preditores melhoram a acurácia da previsão do prêmio de risco por uma margem estatisticamente significativa. Também encontramos que as notícias aumentam a utilidade e os ganhos financeiros para um investidor com função utilidade média-variância.

Palavras-chave: Incerteza Macroeconômica; Regressão de Texto; Aprendizado de Máquina; Previsão Macroeconômica; Previsão Financeira.

Abstract

This dissertation encompasses three chapters that study Economic Forecasting Models. Below are the abstracts for each chapter.

Chapter 1: **Measuring Macroeconomic Uncertainty to Brazil**

The chapter proposes estimating a macroeconomic measure of uncertainty to Brazil. The index was constructed based on the methodology of Jurado (2015) used to build the same index for the US economy. We show that an increase in the uncertainty level leads economic recessions. Moreover, the recent macroeconomic policy adopted by the Brazilian government in 2010-2011 was followed up by substantial increase in the uncertainty level of the Brazilian economy. Our results suggest that the proposed uncertainty measure can be used to assess macroeconomic policies as well as predict economic recessions.

Chapter 2: **Identifying the Predictive Power of FED Minutes**

This chapter proposes a novel method to extract the most predictive information from FED minutes. Instead of considering a dictionary (set of words) with a fixed content, we construct a dictionary whose content is allowed to change over time. Specifically, we utilize machine learning to identify the most predictive words (the most predictive content) of a given minute and use them to derive new predictors. We show that the new predictors improve forecast accuracy of Output growth by a statistically significant margin, suggesting that the combination of machine learning and text regression can be interpreted as a powerful device for out-of-sample macroeconomic forecasting.

Chapter 3: **Equity Premium Forecasting: Identifying the Predictive Power of Financial News**

This chapter proposes a novel method to extract the most predictive information from Financial News published in the Wall Street Journal and The New York Times. Instead of considering a dictionary (set of words) with a fixed content, we construct a dictionary whose content is allowed to change over time. Specifically, we utilize machine learning to identify the most predictive words (the most predictive content) of a given financial news and use them to derive new predictors. We show that the new predictors improve forecast accuracy of Equity Premium by a statistically significant margin. We also find that the Financial News increases the utility and financial gains, for an investor with a mean-variance utility function.

Keywords: Macroeconomic uncertainty; Text Regression; Machine Learning; Macroeconomic Forecasting; Financial Forecasting.

Sumário

List of Figures	9
List of Tables	10
Abreviações	11
1 Medindo Incerteza Macroeconômica para o Brasil	15
1.1 Introdução	15
1.2 Revisão da Literatura	16
1.3 Modelo Econométrico	20
1.4 Dados	21
1.5 Resultados	22
1.5.1 Estimativas da Incerteza Macroeconômica	22
1.5.2 Precedência temporal entre a incerteza e variáveis macroeconômicas	23
1.5.3 Relação entre a volatilidade do IBOVESPA e o índice de incerteza	25
1.5.4 Índice de Incerteza Macroeconômica Agregado com pesos baseados na Análise de Componentes Principais	26
1.5.5 Índice de Incerteza x Leading Indicators	28
1.6 Considerações Finais	29
2 Identificando o Poder Preditivo das Minutas do FED	31
2.1 Introdução	31
2.2 Decisões do Banco Central, curva de juros e crescimento do produto . . .	36
2.3 Text Regression	38
2.3.1 Dados Textuais	38
2.3.2 Construindo Dicionários que variam no tempo	40
2.4 Preditores baseados nas Minutas do FED	42
2.4.1 Dicionários variantes no tempo	42
2.4.2 Dicionários fixos	44
2.4.3 Outros Preditores	45
2.5 Análise Empírica	46
2.5.1 Dados	46
2.5.2 O procedimento de previsão	47
2.5.3 Avaliação da Previsão	50
2.6 Resultados	51
2.6.1 Explicando o sucesso das previsões baseadas nas minutas do FED .	53

2.6.2	Previsões do Livro verde do FED e previsões baseadas nas minutas do FED	55
2.6.3	Análise da Função Impulso Resposta	56
2.7	Considerações Finais	57
3	Previsão do Prêmio de Risco: Identificando o Poder Preditivo das Notícias Financeiras	59
3.1	Introdução	59
3.2	Dados, procedimentos e avaliação da previsão	63
3.2.1	Dados	63
3.2.2	Construindo dicionários variantes no tempo	64
3.2.3	Procedimento de Previsão	65
3.2.4	Avaliação da Previsão	66
3.3	Resultados Empíricos	68
3.3.1	Resultados fora da amostra	68
3.3.2	Explicando os benefícios dos modelos baseados em notícias financeiras	70
3.4	Considerações Finais	72
	 Referências Bibliográficas	 75
	 Apêndice A - Medindo Incerteza Macroeconômica para o Brasil	 81
	 Apêndice B - Identificando o Poder Preditivo das Minutas do FED.	 96
	 Apêndice C - Previsão do Prêmio de Risco: Identificando o Poder Preditivo das Notícias Financeiras.	 111

Lista de Figuras

1	Incerteza Agregada $\overline{\mathcal{U}}_t^y$ para $h=1,3,12$	88
2	Média Móvel 12 meses para a incerteza agregada $h=1,3,12$	88
3	Resposta da produção industrial e do emprego a choques na volatilidade do IBOVESPA	89
4	Volatilidade do IBOVESPA e Medida de Incerteza Macroeconômica.	90
5	Impulso Resposta Incerteza $h=1$ e Volatilidade do IBOVESPA	90
6	Impulso Resposta Incerteza $h=3$ e Volatilidade do IBOVESPA	91
7	Impulso Resposta Incerteza $h=12$ e Volatilidade do IBOVESPA	91
8	Fator Macroeconômico construído a partir do conjunto de variáveis	91
9	Índice de Incerteza Macroeconômica Agregado construído a partir dos pesos obtidos via análise de componentes principais.	92
10	Resposta da Produção Industrial e do Emprego a um choque na Incerteza $h=1$	93
11	Resposta da Produção Industrial e do Emprego a um choque no Índice de Incerteza com pesos baseados nos fatores $h=3$	94
12	Resposta da Produção Industrial e do Emprego a um choque no Índice de Incerteza com pesos baseados nos fatores $h=3$	95
13	Gráfico da série de tempo “economic activity”, 1958(T1)-2017(T2).	96
14	Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.	104
15	Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.	105
16	Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.	106
17	Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.	107
18	Resposta do Yield Spread a um choque no fator Hawkish.	109
19	Resposta do the Yield Spread a um choque no fator Dovish.	110
20	Gráfico de dispersão da variância da previsão e do viés ao quadrado relativo a média histórica, 1999:1–2017:12.	113
21	Riqueza acumulada associada a cada modelo de previsão 1999:1–2017:12.	114

Lista de Tabelas

1	Séries Macroeconômicas Utilizadas na Construção do Índice de Incerteza	81
2	Impacto da Incerteza Individual na incerteza agregada	82
3	Teste de Causalidade de Granger Incerteza x Produção Industrial	83
4	Teste de Causalidade de Granger Incerteza x Emprego	83
5	Decomposição da Variância para a produção industrial, emprego e horas	84
6	Peso de cada série macroeconômica na incerteza agregada	85
7	Teste de Causalidade para o índice de incerteza com pesos baseados em fatores	86
8	Turning Point Produção Industrial e Índice de Incerteza	87
9	QPS(1) Score	87
10	Data de Divulgação do Produto e das Minutas do FED	97
11	RMSFEs para $h = 1, 3, 6$ durante todo o período fora da amostra. Relativo ao Benchmark $AR(2)$	100
12	RMSFEs para $h = 1, 3, 6$ durante os sub períodos de recessão e expansão. Total e relativo ao Benchmark $AR(2)$	101
13	RMSFEs para $h = 0$ durante todo o período fora da amostra. Relativo ao Benchmark $AR(2)$	102
14	RMSFEs para $h = 0$ durante os sub períodos de expansão e recessão. Total and relativo ao Benchmark $AR(2)$	103
15	Termos mais Seleccionados.	108
16	Previsões do Green Book do FED e dos preditores baseados nas minutas do FED	108
17	Resultado fora da amostra para a previsão do Prêmio de Risco.	111
18	Resultado fora da amostra para a previsão do Prêmio de Risco para os sub períodos de expansão e recessão.	112
19	Palavras/termos mais preditivos.	115

Abreviações

PCA	P rincipal C omponent A nalysis
DTM	D ocument T erm M atrix
FED	F ederal Reserve Bank
LASSO	L east A bsolute S hrinkage and S election O perator
FOMC	F ederal O pen M arket C ommittee
LDA	L atent D irichlet A llocation
CODACE	C omitê de D atação de C iclos E conômicos
IBGE	I nstituto B rasileiro de G eografia e E statística
IBOVESPA	Í ndice da B olsa de V alores de S ão P aulo
FRED	F ederal R eserve E conomic D ata
MQO	M ínimos Q uadrados O rdinários
TF-IDF	T erm F requency - I nverse D ocument F requency
NBER	T he N ational B ureau of E conomic R esearch

Dedico esta tese a toda minha família, em especial a minha esposa Beônia, ao meu filho Jailson Godeiro Neto, a minha mãe Fátima, a minha avó Adelziva(in memoriam) e ao meu pai Jailson Godeiro(in memoriam).

Introdução

A previsão de variáveis econômicas é um dos maiores desafios para os Economistas, principalmente nos últimos anos, com o surgimento de novas técnicas e dados. Atualmente, temos a influência do Big Data, onde a cada minuto são produzidos milhões de informações, sejam elas via publicações em redes sociais, scanner data (dados dos consumidores ao efetuarem suas compras), entre outros. O uso de dados não estruturados vem ganhando importância nos últimos anos, com o avanço da capacidade de processamento e armazenamento de informações por parte dos computadores. [Gentzkow et al. \(2017\)](#) afirma que os dados não estruturados são um rico complemento aos preditores econômicos estruturados. Nesse contexto, torna-se essencial não apenas usar os tradicionais preditores econômicos disponíveis, mas também a busca de novos preditores. Nesse sentido, o presente trabalho investiga novos métodos e técnicas que buscam melhorar o desempenho de modelos de previsão econômica.

A tese é composta por três capítulos. No capítulo [1](#), usamos indicadores antecedentes para construir um índice de incerteza para o Brasil, baseado em [Jurado et al. \(2015\)](#). Vale ressaltar que uma das principais motivações desse primeiro capítulo deve-se ao Brasil ter enfrentado a maior recessão da sua história durante os anos de 2014 e 2016. Então, ao verificarmos que [Jurado et al. \(2015\)](#) construiu um índice de Incerteza Macroeconômica para os Estados Unidos, objetivamos por meio do capítulo 1 usar essa mesma metodologia na construção de um índice de Incerteza para a economia brasileira. Neste capítulo uma das principais conclusões é que o índice criado supera a volatilidade do mercado de ações, proxy mais conhecida da incerteza.

No capítulo [2](#), usamos informações das minutas do FED para prever o crescimento do Produto dos Estados Unidos, usando métodos de Aprendizado de Máquina Supervisionado para selecionar as palavras mais preditivas em um dado período do tempo. Uma das motivações do capítulo deve-se ao fato de que, a partir da crise de 2008, surge o que chamamos de “unconventional monetary policy”, onde as taxas de juros chegam a zero, e o FED começa a adotar novas políticas como o quantitative easing, a fim de controlar os juros de longo prazo principalmente. Nesse cenário, a comunicação do FED ganha uma

importância ainda maior, pois por meio dela o Banco Central pode controlar as expectativas e consequentemente os juros de longo prazo. Também o que nos motiva é o fato de que, como dito na citação do primeiro parágrafo, os dados não estruturados podem ser um rico complemento aos dados tradicionais. Mostramos que o poder preditivo das minutas do FED é não trivial, dado que apenas o dicionário variável tem poder preditivo.

No capítulo 3, usamos as Notícias financeiras publicadas no “The Wall Street Journal” e no “The New York Times”, a fim de verificar se as mesmas tem poder preditivo e são capazes de incrementar o retorno dos agentes financeiros. Trabalhos como os de [Garcia \(2013\)](#) estudam esse assunto, mas utiliza um dicionário fixo e não realiza um exercício empírico fora da amostra. Outra motivação é que alguns fundos de investimento na atualidade, definem a alocação de seu portfólio baseado não no lucro ou fundamentos das empresas, mas com base no sentimento que vem das notícias sobre aquela empresa e/ou post dos investidores sobre aquela empresa em redes sociais como o Twitter¹, por exemplo.². Então, se tais gestores usam notícias objetivando melhorar o desempenho de uma carteira, isto nos indica que as mesmas podem ter alguma importância, por isso o capítulo testa essa hipótese empiricamente. Além do mais, em um paper seminal [Bollen et al. \(2011\)](#) demonstra o poder preditivo das informações provenientes de redes sociais como o twitter. Como principais resultados do capítulo verificamos que as Notícias Financeiras tem um forte poder preditivo, mostrando tanto ganhos econômicos quanto estatísticos para os investidores. Um investidor com utilidade média variância pode potencializar seus ganhos financeiros, ao usar um modelo baseado em notícias ao invés da média histórica e outros modelos tradicionalmente usados.

¹No Twitter por exemplo, ao usarmos o \$ juntamente com o ticker da empresa, isto funciona como uma hashtag. Então, por exemplo, caso queiramos emitir uma opinião ou fazer uma busca sobre a empresa Apple, por exemplo, basta digitarmos \$AAPL.

²O ETF negociado na Bolsa de Nova Iorque, com o ticker [BUZ](#) é um exemplo. O mesmo escolhe suas ações baseado no sentimento das notícias e redes sociais.

Capítulo 1

Medindo Incerteza Macroeconômica para o Brasil

1.1 Introdução

A incerteza é um dos aspectos mais importantes, no que diz respeito ao investimento e ao crescimento econômico (Bachmann et al., 2013). Um cenário de incerteza ocorre quando há dificuldade na previsão de cenários futuros e no planejamento por parte das famílias e das empresas. Em um ambiente de incerteza, as famílias tendem a consumir menos e as empresas a investirem menos, levando a um menor crescimento econômico ou até mesmo a recessão.

O momento econômico pelo qual o Brasil atravessou nos anos de 2014 e 2015 caracterizou-se pelo aumento da incerteza macroeconômica, levando os agentes a uma maior dificuldade de prever as principais variáveis para o seu processo de decisão. Há um debate a respeito dos efeitos das políticas iniciadas a partir de 2008 e aprofundadas em 2011, com a chamada nova matriz macroeconômica.¹ O debate econômico sugere que tais medidas podem ter contribuído para aumentar o grau de incerteza macroeconômica, mas nenhum estudo acadêmico foi realizado no sentido de identificar essa possibilidade. Portanto, a primeira questão a ser respondida por este artigo é se a nova matriz macroeconômica aumentou de fato o nível de incerteza econômica no Brasil, contribuindo para o surgimento da recessão recente ocorrida entre os anos de 2014 e 2016, de acordo com o CODACE.

¹O conjunto de medidas introduzido no Brasil como uma resposta para a crise econômica mundial ficou conhecido como “a nova matriz macroeconômica” cuja definição pode ser encontrada neste link: <http://www.ptnosenado.org.br/site/noticias/ultimas/item/37758-nova-matriz-economica-garantira-crescimento-acelerado>.

O índice de incerteza Macroeconômica agregada para o Brasil será construído baseado na metodologia desenvolvida por [Jurado et al. \(2015\)](#) e que também foi usada para calcular o índice de incerteza da economia americana. Nesse sentido, a segunda contribuição deste artigo é de verificar se o índice de incerteza proposto consegue capturar a incerteza melhor do que outras medidas utilizadas na literatura, tal como a volatilidade do mercado de ações. Para a construção do índice usamos as mesmas séries macroeconômicas de [Issler et al. \(2013\)](#), que constituem séries consideradas como indicadores antecedentes da atividade econômica. A tabela 1 mostra todas as séries utilizadas. Os nossos resultados apontam para uma correlação negativa de -0.45 entre o índice de incerteza macroeconômica e variáveis reais, tais como a produção industrial. Testes de causalidade de Granger sugerem que incerteza causa produção industrial, mas a produção industrial não causa incerteza. Por fim, encontramos que a nova matriz macroeconômica contribuiu para o aumento da incerteza econômica no Brasil levando, via causalidade de Granger, a uma queda da produção industrial. Por fim, nossos resultados sugerem que o índice proposto de incerteza pode ser usado para prever recessões econômicas, constituindo assim a terceira contribuição deste artigo. Além da introdução e das considerações finais, a seção 1.2 apresenta a revisão da literatura sobre incerteza, a seção 1.3 trata do modelo econométrico, a seção 1.4 dos dados utilizados e a seção 1.5 discute os resultados encontrados pela pesquisa.

1.2 Revisão da Literatura

A discussão sobre como medir incerteza é recente, tendo como principais trabalhos [Jurado et al. \(2015\)](#), [Bloom \(2009\)](#), [Carriero et al. \(2016\)](#). Com efeito, [Jurado et al. \(2015\)](#) constroem uma medida de incerteza para os Estados Unidos e comparam sua eficácia em captar as oscilações nos ciclos de negócios com proxies para incerteza geralmente utilizadas, como, por exemplo, a volatilidade. A medida de incerteza de [Jurado et al. \(2015\)](#) teve uma correlação negativa maior com a produção industrial do que a volatilidade, o que significa que a medida de incerteza construída pelos autores conseguiu expressar mais fortemente as oscilações das principais variáveis da economia. Por exemplo, a correlação do índice de incerteza 1 passo a frente com a produção industrial americana foi -0.59, ao passo que a correlação da produção industrial com a volatilidade do mercado de ações foi de apenas -0.30. Esse resultado indica que a medida obtida por [Jurado et al. \(2015\)](#) é mais eficaz do que a volatilidade, no sentido de antecipar recessões econômicas.

Jurado et al. (2015) define a incerteza h períodos a frente na variável $y_{jt} \in Y_t = (y_{1t}, \dots, y_{N_{y_t}})'$ denotado por $\mathcal{U}_{jt}^y(h)$ como sendo a volatilidade condicional do componente imprevisível do valor futuro da série. Matematicamente, temos:

$$\mathcal{U}_{jt}^y(h) \equiv \sqrt{E \left[(y_{jt+h} - E[y_{jt+h}|I_t])^2 | I_t \right]} \quad (1.1)$$

Onde a esperança $E(\cdot|I_t)$ é tomada com respeito a informação disponível em I_t aos agentes econômicos. De acordo com Jurado et al. (2015), se hoje a esperança (condicionada a toda a informação disponível) do erro quadrado da previsão de y_{jt+h} aumenta, então a incerteza em y_{jt} aumentará. Uma medida ou índice de incerteza macroeconômica pode ser construído agregando as incertezas individuais usando os pesos w_j :

$$\mathcal{U}_t^y \equiv p \lim_{N_y \rightarrow \infty} \sum_{j=1}^{N_y} w_j \mathcal{U}_{jt}^y(h) \equiv E_w[\mathcal{U}_{jt}^y(h)]. \quad (1.2)$$

Assim, de acordo com a metodologia de Jurado et al. (2015), a construção do índice de incerteza segue três passos:

1. Obter a previsão das N séries macroeconômicas $E[y_{jt+h}|I_t]$. Pode-se utilizar análise fatorial e construir fatores.
2. Definindo o horizonte de previsão h , obter o erro de previsão $V_{jt+h}^y \equiv y_{jt+h} - E[y_{jt+h}|I_t]$, e estimar sua variância, $E[(V_{jt+h}^y)^2|I_t]$.
3. Estimar o fator de incerteza macroeconômica $\mathcal{U}_t^y(h)$, agregando as incertezas individuais $\mathcal{U}_{jt}^y(h)$. Pode-se utilizar a ponderação igual para cada fator individual $1/N$ ou ainda outros métodos de estimação dos pesos.

Bloom (2009) estuda os choques da incerteza na economia por meio do modelo VAR, onde permite-se que o segundo momento varie no tempo. Os choques na incerteza produzem, segundo Bloom (2009), uma queda em variáveis macroeconômicas reais, tais como o produto e o emprego. De acordo com o autor, isso ocorre porque um alto índice de incerteza provoca uma diminuição no investimento das empresas e, conseqüentemente, uma retração na contratação de trabalhadores.

Jurado et al. (2015) destacam a importância de se utilizar a medida de incerteza ao invés de variáveis como a volatilidade do mercado de ações por exemplo. Segundo os autores, episódios de incerteza ocorrem com menos frequência do que indicado pelos

índices de volatilidade, mas quando eles ocorrem mostram maior correlação e persistência com o índice de incerteza. Além disso, períodos de recessão prolongada são associados a aumentos no índice de incerteza. A presente pesquisa identificará se os resultados encontrados para a economia americana se sustentam para a economia brasileira.

Segundo [Bachmann et al. \(2013\)](#), existe uma interação entre alta incerteza e baixo crescimento econômico ou até mesmo recessões, o que provoca pausa nas contratações e no investimento - as firmas “esperam e veem” (wait and see) como o futuro se revela. Isso causa uma queda na atividade econômica. Depois de um certo número de períodos, há uma demanda reprimida por fatores de produção, tal que a queda inicial é seguida por uma rápida recuperação da atividade econômica.

[Scotti \(2016\)](#) constrói um índice de incerteza para os Estados Unidos, Zona do Euro, Reino Unido, Canadá e Japão baseado no erro de previsão dos analistas financeiros coletados pela Bloomberg para os dados macroeconômicos. O estudo toma como base a mediana da previsão dos analistas financeiros. Ainda segundo [Scotti \(2016\)](#), o erro de previsão também é conhecido como uma “notícia surpresa” ou simplesmente “surpresa”. Para estimar os pesos de cada variável macroeconômica no índice, a autora usa o modelo de fatores, onde supõe-se que os agentes usam toda a informação disponível na previsão e o erro de previsão, como dito anteriormente, representa a surpresa. Assim, ainda segundo [Scotti \(2016\)](#) a eficiência da previsão pode ser testada a partir da seguinte equação:

$$s_t^i = \alpha^i + \beta^i y_t^{i,f} + u_i^t \quad (1.3)$$

Onde $s_t^i = y_t^i - y_t^{i,f}$ é o erro de previsão. A hipótese nula do teste é $\alpha^i = \beta^i = 0$. Se a hipótese nula for aceita, então a previsão não tem correlação com o erro de previsão e, além disso, o valor esperado do erro de previsão é 0. De acordo com o teste feito por [Scotti \(2016\)](#), a previsão dos analistas financeiros divulgados pela Bloomberg não foi eficiente na maioria dos casos, dado que α^i e β^i foram diferentes de zero para a maioria das séries pesquisadas. Mesmo não sendo eficiente a previsão divulgada pela Bloomberg, [Scotti \(2016\)](#) justifica o seu uso na pesquisa, dado que a mesma é publicamente disponível e também dado que o mercado reage a previsão que a Bloomberg divulga, e não a uma suposta previsão construída pela autora. A partir do erro de previsão computado na equação 1.3, [Scotti \(2016\)](#) calcula o índice de incerteza, a partir da equação:

$$u_t = \sqrt{\sum_{j=1}^t w_j s_j} \quad (1.4)$$

A diferença da metodologia aplicada por [Scotti \(2016\)](#) e a metodologia empregada pela presente pesquisa e por [Jurado et al. \(2015\)](#) é que a variável s_j representa o erro de previsão com base nas estimativas dos analistas da Bloomberg, e não com base na previsão de um modelo econométrico. E como já demonstrado por [Scotti \(2016\)](#), o mesmo não performa bem nos testes de eficiência. No entanto, pela definição de incerteza de [Jurado et al. \(2015\)](#), precisa-se retirar toda a informação disponível da série para o erro representar a incerteza. No entanto, o foco da pesquisa de [Scotti \(2016\)](#) não é apenas a incerteza em si, mas também o impacto da surpresa nas variáveis reais da economia.

[Bloom et al. \(2012\)](#) analisam os choques de incerteza como sendo um novo choque que direciona os ciclos econômicos. Os autores demonstram que a incerteza macroeconômica é contracíclica e aumenta agudamente em recessões, especialmente na grande recessão de 2008. Em simulações feitas pelos autores utilizando o modelo DSGE, nota-se que um aumento na incerteza reduz o crescimento da produtividade, devido a redução no grau de realocação na economia. Ainda segundo [Bloom et al. \(2012\)](#) alta incerteza faz com que plantas produtivas pausem sua expansão e fábricas improdutivas pausem seus contratos.

Outra questão importante a ser analisada segundo [Bloom et al. \(2012\)](#) é a relação de causalidade entre incerteza e variáveis macroeconômicas, como produção industrial, por exemplo. Será que um aumento da incerteza acarreta uma queda na produção industrial, ou seria um aumento na incerteza gerada endogenamente dado queda nas variáveis macroeconômicas. [Bloom et al. \(2012\)](#) argumenta que a literatura recente sugere alguns mecanismos que fazem com que a incerteza aumente endogenamente em épocas de recessão.

A pesquisa de [So \(2013\)](#) indica que os preços não refletem corretamente os componentes previsíveis, dado que existe viés na previsão dos analistas. Com isso, os investidores alocando os recursos de seus investimentos com base na previsão dos analistas, tendem a apreçar de forma errada os ativos. Essa é uma das críticas aos trabalhos que calculam a incerteza com base na previsão dos analistas. Esse também foi um dos argumentos que [Jurado et al. \(2015\)](#) justifica o uso de uma previsão feita por um modelo econométrico no cálculo do índice de incerteza. Além do mais, segundo [Jurado et al. \(2015\)](#) das 132 séries macroeconômicas pesquisadas para o seu trabalho, apenas um quinto eram acompanhadas pelos analistas de mercado.

1.3 Modelo Econométrico

Seja a série y_{jt+h} a série para ser calculada, a incerteza no período $h \geq 1$. De acordo com a metodologia apresentada por [Jurado et al. \(2015\)](#), o econometrista precisa escolher um modelo de previsão para y_{jt+h} . Neste artigo, utilizamos um modelo de Vetores auto regressivos (VAR), desenvolvido por [Sargent et al. \(1977\)](#), para calcular a previsão h -passos a frente das variáveis y_{jt+h} . Em outras palavras, seja \mathbf{Y}_{jt} um vetor com as variáveis macroeconômicas que se deseja realizar a previsão h -passos a frente. Assim, temos um modelo VAR de forma reduzida representado pela equação:

$$\mathbf{Y}_{jt} = \Phi_j^Y \mathbf{Y}_{jt-1} + \mathbf{V}_{jt}^Y. \quad (1.5)$$

Sob uma função perda quadrática e assumindo que as variáveis são estacionárias, a previsão ótima h passos a frente será:

$$E\mathbf{Y}_{jt+h} = (\Phi_j^Y)^h \mathbf{Y}_{jt}. \quad (1.6)$$

A variância do erro de previsão no tempo t será:

$$\Omega_{jt}^Y(h) = E_t[(\mathbf{Y}_{jt+h} - E_t\mathbf{Y}_{jt+h})(\mathbf{Y}_{jt+h} - E_t\mathbf{Y}_{jt+h})']. \quad (1.7)$$

Para $h=1$ temos:

$$\Omega_{jt}^Y(1) = E_t(\mathbf{V}_{jt+1}^Y \mathbf{V}_{jt+1}^{Y'}). \quad (1.8)$$

Para $h > 1$, a variância do erro de previsão de \mathbf{Y}_{jt+h} evolui de acordo com:

$$\Omega_{jt}^Y(h) = \Phi_j^Y \Omega_{jt}^Y(h-1) \Phi_j^{Y'} + E_t(\mathbf{V}_{jt+1}^Y \mathbf{V}_{jt+1}^{Y'}). \quad (1.9)$$

Ou seja, de acordo com [Jurado et al. \(2015\)](#) quando $h \rightarrow \infty$, a previsão tende para a média incondicional e variância do erro de previsão é a variância não condicional de \mathbf{Y}_{jt+h} .

Assim, pode-se computar a incerteza de cada série a partir da equação:

$$\mathcal{U}_{jt}^y(h) = \sqrt{\mathbf{1}_j' \Omega_{jt}^Y \mathbf{1}_j}. \quad (1.10)$$

A partir da incerteza individual de cada série calcula-se a incerteza agregada:

$$\mathcal{U}_t^h(h) = \sum_{j=1}^{N_y} w_j \mathcal{U}_{jt}^y(h). \quad (1.11)$$

Utilizamos além dos pesos iguais para cada incerteza individual $1/N_y$, os pesos baseados nos loadings do primeiro componente principal calculado a partir das séries macroeconômicas apresentadas na tabela 1.

1.4 Dados

As séries macroeconômicas utilizadas na construção da medida de incerteza foram baseadas nas séries usadas por Issler et al. (2013) na construção de índices coincidentes e leading indicators para a economia brasileira. A tabela 1 apresenta as séries utilizadas. Os dados foram coletados no site do Ipeadata, do Banco Central do Brasil e do Instituto Brasileiro de Geografia e Estatística (IBGE). O período pesquisado inicia-se em Janeiro de 2002 e vai até Junho de 2015. A frequência utilizada foi mensal.

Os dados nominais foram deflacionados pelo índice de preços - Disponibilidade Interna (IGP-DI) calculado pela Fundação Getúlio Vargas (FGV). A base utilizada para o deflacionamento foi junho de 2015, assim, todas as séries estão a preços de 06/2015. A partir da análise do correlograma das séries, foi identificada a presença ou não de sazonalidade. Para as séries com sazonalidade foi efetuado o ajustamento sazonal via Filtro de Kalman. O procedimento realizado consiste em dois passos. Primeiro identifica-se se a série segue o modelo de nível ou tendência local com sazonalidade. Em seguida estima-se o modelo e retira-se o efeito sazonal captado. Assim, obtêm-se a nova série sem a presença de sazonalidade.

[Tabela 1 aqui]

Em seguida, foram testados estacionaridade para todas as séries. Foi utilizado o teste ADF com constante e tendência para verificar estacionaridade das séries. As séries $I(1)$ foram diferenciadas. Para a previsão foi utilizado o modelo vetor autor regressivo de ordem 1 VAR(1). O modelo VAR(1) foi escolhido utilizando-se o critério de informação BIC, dado que o mesmo prioriza modelos mais parcimoniosos. Os horizontes escolhidos foram $h = 1, 3$ e 12 . Foi utilizado o método de previsão recursivo, iniciando com as 60 observações iniciais, dado que usaremos o método de previsão recursivo e a nossa amostra está sendo estendida a cada instante de tempo t . Então considerando $h = 1$, por

exemplo, a partir da estimação dos parâmetros utilizando como amostra as 60 primeiras observações, foi realizado a previsão um passo a frente para $t=61$. Em seguida, o modelo foi re-estimado incluindo a observação 61 e partir desse modelo previsto até $t = 62$, e assim sucessivamente. Assim, foram obtidos 100 previsões, que abrange o período de março de 2007 até junho de 2015. Os resultados seguem na próxima sessão.

1.5 Resultados

1.5.1 Estimativas da Incerteza Macroeconômica

A figura 1 apresenta as estimativas para a incerteza macroeconômica para os horizontes $h=1,3,12$. Em linha com os resultados de [Jurado et al. \(2015\)](#), a incerteza macroeconômica para o Brasil é contracíclica e apresenta correlação negativa com a produção industrial, -0.19, -0.32, -0.33, para $h=1,3,12$. Em outras palavras, um aumento do índice de incerteza indica queda da produção industrial.

[Figura 1 aqui]

[Figura 2 aqui]

O índice consegue, em alguns períodos, captar indícios de recessão na economia. Por exemplo, a partir do início de 2008 já percebe-se um aumento no índice, o que seria um indício da recessão que ocorreu em 2009. Outro aumento do índice ocorre a partir de 2013, indicando a recessão prolongada que o país viria a enfrentar a partir de 2014. Outro fato observado no gráfico é o aumento da incerteza macroeconômica já a partir do final de 2011. Esse período coincide com o início da quebra do tripé macroeconômico no Brasil, em que o Banco Central inicia um processo de queda na taxa de juros mesmo sem a inflação estar no centro da meta, e o governo federal intensifica as políticas populistas como controle de preços, câmbio, e empréstimos subsidiados via BNDES. Esse conjunto de medidas “heterodoxas” recebeu o nome de “A nova matriz macroeconômica” mas no fundo representou um abandono do tripé econômico baseado na geração de superavit fiscal, perseguição da meta de inflação e câmbio flutuante. Analisando a figura 1 percebe-se claramente que o índice de incerteza em 2010 era baixo, menos de 1. Como dito anteriormente, com a implantação do que foi chamado “A nova matriz macroeconômica”, há um aumento claro no índice. Com o aumento da incerteza, a confiança do empresário tende a diminuir e o investimento também, contribuindo para o surgimento de recessões econômicas, como aquela iniciada em 2014.

A mudança no nível de incerteza correlaciona-se com a variação no PIB do Brasil. Com efeito, a incerteza sobe em 2009 e o PIB cai 0.2%, enquanto que em 2010 observa-se uma queda no índice e um aumento do PIB de 7.6%. Assim, verifica-se que quanto maior a taxa de crescimento do PIB menor tende a ser a incerteza. A tabela 2 mostra o impacto da incerteza individual na incerteza agregada. De forma análoga a [Jurado et al. \(2015\)](#), a pesquisa tenta identificar o quanto cada fator de incerteza individual explica a incerteza macroeconômica total. O procedimento efetuado consiste em uma regressão por MQO de \mathcal{U}_{jt}^y sobre $\overline{\mathcal{U}}_t^y$ e o respectivo cálculo do R^2 . Para $h=1$, nota-se que a produção de papelão tem um maior poder de explicação sobre a incerteza agregada, com um R^2 de 0.19. Já para $h=3$, destaca-se o IBOVESPA, com um R^2 de 0.33. Para $h=12$, temos a produção industrial com maior poder de explicação sobre o fator agregado, como um R^2 de 0.15. Portanto, não existe um fator que se destaca na explicação do fator agregado, dado que o maior valor de R^2 foi 0.33. Então, observa-se que são várias as fontes que contribuem para a formação do índice de incerteza macroeconômica, e não apenas uma fonte em especial.

[Tabela 2 aqui]

1.5.2 Precedência temporal entre a incerteza e variáveis macroeconômicas

O primeiro passo é identificar uma possível precedência temporal (causalidade de Granger) entre a produção industrial, o emprego e a incerteza. As tabelas 3 e 4 apresentam os testes de causalidade de Granger entre os índices de incerteza com $h=1,3,12$ e variáveis macroeconômicas como a produção industrial e o emprego. As defasagens para a estimação do VAR foram escolhidos de acordo com o critério de informação de Schwarz, dado que o mesmo prioriza o modelo mais parcimonioso. Foi estimado o VAR Bi-Variado, entre incerteza e produção industrial e entre incerteza e emprego. As variáveis foram estimadas em nível para a incerteza e na primeira diferença para a produção industrial e o emprego. O lag de maior ordem escolhido foi 3, para a relação do índice de incerteza $\overline{\mathcal{U}}(12)$ com a produção industrial e o emprego respectivamente. Para o VAR que incluiu $\overline{\mathcal{U}}(1)$ e $\overline{\mathcal{U}}(3)$ o lag escolhido foi de ordem 2.

[Tabela 3 aqui]

[Tabela 4 aqui]

Para $\bar{U}(1)$, constata-se que ao nível de significância de 5% que a incerteza não Granger-cause a produção industrial. No entanto, para $\bar{U}(3)$ e $\bar{U}(12)$ a 5% constata-se que a incerteza granger-cause a produção industrial, mas a produção industrial não granger causa a incerteza. Em outras palavras, nossos resultados apontam que incerteza ajuda a prever produção industrial, mas produção industrial não ajuda a prever incerteza. Já com relação ao emprego, o teste de causalidade de Granger indica que $\bar{U}(1)$ causa o emprego no presente exercício, mas o emprego não Granger causa $\bar{U}(1)$. Os resultados sugerem que o índice de incerteza é um bom indicador na previsão e antecipação de uma possível recessão econômica, dado que ela Granger-cause (ajuda a prever) mudanças em variáveis macroeconômicas reais tais como a produção industrial e o emprego. Assim, de posse do índice, formuladores de política econômica podem antecipar uma possível queda na produção industrial resultante de um aumento da incerteza econômica calculada pelo índice ora proposto.

Portanto, além do índice de incerteza apresentar uma correlação negativa com a produção industrial, o mesmo também precede no tempo variáveis macroeconômicas reais, como produção e o emprego, dado que incerteza causa no sentido de Granger a produção industrial e o emprego. Em um cenário de aumento de incerteza, onde os agentes econômicos passam a ter uma menor previsibilidade da economia, empresas e consumidores adiam decisões de consumo e investimento. Este é a teoria conhecida por “wait and see”, abordada por [Bloom et al. \(2012\)](#). Diferente do trabalho de [Bloom et al. \(2012\)](#), onde segundo a literatura recente a incerteza era gerada endogenamente pela queda das variáveis macroeconômicas reais, para o caso brasileiro claramente há uma relação de causalidade. Para o Brasil, um aumento de incerteza provoca queda na produção e no emprego, como demonstrado nos testes. Ao analisarmos a figura 2, nota-se que em 2008, por exemplo, já ocorre um aumento na incerteza. A mesma girava em torno de 1.05%, e a partir de meados de 2008 aumenta para próximo de 1.3%. No entanto a recessão ocorre em 2009 apenas. Para o ano de 2013, também notamos que a incerteza que estava abaixo de 1%, começa a aumentar a partir do segundo semestre de 2013, sendo que a alta continua em 2014 e 2015. Outro fato interessante é que em 2013 ocorre uma inflexão na série da incerteza, onde a mesma vinha decrescente, depois volta a aumentar. Essa inflexão já representa indícios da recessão que viria a ocorrer nos anos de 2014, 2015 e 2016.

Esse resultado, que incerteza causa produção industrial e emprego para o Brasil, sugere que políticas econômicas erradas aumentam a incerteza no ambiente de negócios e reduzem a produção do país. Desta forma, há indícios de que as políticas econômicas iniciadas a partir de 2008 e aprofundadas em 2011 aumentaram a incerteza do país, e com isso tiveram uma contribuição elevada na determinação da recessão econômica que afeta o Brasil. Ou seja, ao invés da nova matriz macroeconômica incentivar a economia

ela a desestimulou, principalmente porque o mercado percebeu que as políticas adotadas de cunho populista estavam na direção errada, aumentando a incerteza e reduzindo a produção. Assim, políticas que mantenham a estabilidade no ambiente de negócios e reduzam a incerteza são importantes para o crescimento do país. Em um ambiente estável, com inflação controlada e onde os tomadores de decisões conseguem de forma não tão sofisticada muitas vezes prever o futuro, há incentivos ao investimento e a contratação de novos trabalhadores, gerando portanto crescimento econômico e prosperidade.

1.5.3 Relação entre a volatilidade do IBOVESPA e o índice de incerteza

Outra a análise a ser feita é a relação entre o índice de incerteza macroeconômica e a volatilidade do mercado de ações, dado que a volatilidade muitas vezes é utilizada como proxy para incerteza, como em [Bloom \(2009\)](#). Para analisar a relação entre a incerteza e a volatilidade do mercado de ações, [Jurado et al. \(2015\)](#) utilizaram o índice de volatilidade VXO, que é um índice de volatilidade implícita calculado pelo “Chicago Board of Options Exchange”, a partir dos preços dos contratos de opções do índice S&P 100. Como para o Brasil não dispomos de índices de volatilidade implícita, a estratégia escolhida pela pesquisa foi estimação da volatilidade para o mercado de ações brasileiro via modelo GARCH(1,1) desenvolvido por [Engle and Bollerslev \(1986\)](#). A escolha do modelo GARCH (1,1) se deu pela sua excelente performance na previsão de volatilidade como documentada na literatura (i.e., [Gaglianone et al. \(2012\)](#)). O GARCH(1,1) também serve para aproximar um $ARCH(\infty)$ e assim é considerado um modelo parcimonioso. Outros modelos de volatilidade poderiam ser usados, mas não há certeza de que eles teriam um poder de previsão superior ao GARCH (1,1). Além do mais o GARCH foi usado apenas para gerar a volatilidade do IBOVESPA, pois a ideia é mostrar como a volatilidade do IBOVESPA, sendo uma medida alternativa de incerteza, se relaciona como a medida proposta.

As medidas para o índice de incerteza e a volatilidade do IBOVESPA foram padronizadas, com base na distribuição normal padrão. A padronização segue [Jurado et al. \(2015\)](#) e consiste em deixar as variáveis na mesma escala, no qual usa-se o procedimento de retirar do valor da variável sua média e dividir pelo desvio padrão.

[Figura 4 aqui]

Verifica-se uma alta correlação entre o índice de incerteza e a volatilidade do IBOVESPA. No entanto, nota-se em alguns momentos onde há aumento no índice de incerteza muitas vezes não acompanhado por aumentos na volatilidade. Um desses períodos é o

início de 2014, onde ocorre um aumento na incerteza, influenciado principalmente pelas eleições e pela desaceleração da economia, no entanto verifica-se um aumento na volatilidade do mercado de ações apenas no final de 2014. Neste ano os movimentos na volatilidade do mercado de ações eram influenciados principalmente pelo cenário eleitoral e pelas pesquisas de intenção de voto. Notícias que indicavam a vitória da candidata Dilma Rousseff causavam queda no mercado de ações, ao passo que notícias de possível vitória do candidato Aécio Neves traziam aumentos nos índices da bolsa.

Seguindo o procedimento de [Jurado et al. \(2015\)](#) usamos o critério de informação de Schwarz para estimar um VAR(12) bi-variado com a variável de incerteza a volatilidade do IBOVESPA. Os resultados seguem nas figuras 5, 6 e 7. Observa-se que aumentos na volatilidade do IBOVESPA aumentam a incerteza, mas aumentos na incerteza não impactam a volatilidade. Os resultados encontrados para os dados brasileiro não estão em linha com os resultados de [Jurado et al. \(2015\)](#). O que pode explicar essa questão é o fato de o mercado brasileiro ser altamente influenciado por mercados internacionais de ações. No entanto essa questão não é escopo dessa pesquisa. Entretanto, a correlação do índice de incerteza e da volatilidade do Ibovespa é de 0.45, correlação idêntica a encontrada por [Jurado et al. \(2015\)](#) para os dados americanos.

Outro fato importante é que a correlação da volatilidade do IBOVESPA com a produção industrial foi de -0.10, uma correlação maior do que a correlação do índice de incerteza com a produção industrial, que foi de -0.33. Com relação a decomposição da variância, nota-se também um maior poder de explicação da medida de incerteza a variáveis como Emprego, Horas e Produção industrial. Por exemplo, o índice de incerteza explica 8.53% da variância do emprego no lag 10, ao passo que a volatilidade do Ibovespa explica apenas 1.97%. A diferença entre os dois modelos é que o primeiro é estimado com o índice de incerteza e o segundo com a volatilidade do IBOVESPA. O índice de incerteza com $h=12$ explica 18.18% da variação das horas de trabalhadas na indústria, enquanto que a volatilidade do IBOVESPA explica 14.56%.

1.5.4 Índice de Incerteza Macroeconômica Agregado com pesos baseados na Análise de Componentes Principais

O índice de incerteza construído nas seções anteriores admitia por simplicidade participações iguais $1/N$ para cada série macroeconômica. No entanto, sabe-se que pode ocorrer que alguma série tenha uma maior influência do que outras no que diz respeito à incerteza. Os pesos de cada série macroeconômica serão obtidos via a análise de componentes principais. Serão utilizados os pesos do fator de maior influência, ou seja, o que representa a maior variabilidade das séries. Os pesos serão normalizados a fim

de somarem 1, para que possamos comparar o índice com pesos a partir da análise de componentes principais com os índices com pesos $1/N$. Os pesos seguem na tabela 6.

[Tabela 6 aqui]

Observa-se que as variáveis macroeconômicas de maior peso foram: termos de troca, exportações e produção industrial. Este resultado indica que uma maior incerteza individual nestas séries tendem a impactar mais o índice agregado. Variáveis como Ibovespa e INPC também tiveram um peso relevante na construção do índice agregado, com pesos normalizados acima de 5%. O fator no qual os pesos foram extraídos explica 45% da variação das séries macroeconômicas. A variável construída a partir da análise de componentes principais segue na figura 8.

[Figura 8 aqui]

Os índices de incerteza para $h=1,3,12$ construídos a partir dos pesos baseados nos fatores seguem na figura 9. Notamos que o índice continua contracíclico com a produção industrial, tendo uma correlação de -0.20, -0.37 e -0.44, para $h=1,3$ e 12, respectivamente. Com a utilização dos pesos obtidos a partir dos fatores, verifica-se que o índice apresenta uma melhora com relação ao índice $1/N$, dado o índice de incerteza com base nos fatores apresenta correlação menor com a produção industrial do que o índice com pesos $1/N$. Por exemplo, para $h=1$, a correlação passa de -0.08 obtida usando o índice $1/N$ para -0.20 no índice com o peso baseado nos fatores. Esse resultado melhora a performance do índice de incerteza, aumentando o seu potencial de antecipação de cenários de recessão econômica. Isto ocorre porque agora é dado um maior pesos a séries que supostamente influenciam mais no aumento da incerteza do que séries com menor influência.

A tabela 7 mostra o teste de causalidade para o índice com pesos baseados em fatores. Os resultados mostram que o índice de incerteza baseado em fatores causa produção industrial e emprego, resultado similar ao observado para o índice com pesos $1/N$. Para este índice baseado em fatores também observamos para $h = 3,12$ que a produção industrial causa incerteza a 5%. Neste sentido o índice precede no tempo estas variáveis macroeconômicas, sugerindo que a incerteza consegue captar variações na produção industrial e no emprego.

[Tabela 7 aqui]

1.5.5 Índice de Incerteza x Leading Indicators

Essa sessão compara os resultados obtidos por [Issler et al. \(2013\)](#) para os leading indicators com os resultados obtidos pelo índice de incerteza proposto neste artigo. Segundo [Issler et al. \(2013\)](#), a medida utilizada para medir a qualidade no qual uma série antecipa o estado da economia é dado pelo *QPS* (*Quality Probability Score*), proposto por [Diebold and Rudebusch \(1990\)](#), isto é:

$$QPS(h) = \frac{\sum_{t=1}^T (P_t - R_t)^2}{T} \quad (1.12)$$

onde P_t denota o estado previsto pelo leading indicator, R_t denota a realização observada na série de referência. Ambos são iguais a 1 para um turning point e zero caso contrário. Assim, $QPS(h)=0$ denota previsão perfeita do estado da economia por parte do índice. Ainda de acordo com [Issler et al. \(2013\)](#), o leading indicator deve causar a série de referência para se efetuar o cálculo do QPS score. Nesse caso, as sessões anteriores mostram que o índice de incerteza causa a produção industrial.

Assim, foram calculados o QPS score para os índices de incerteza com $h=1,3,12$ tendo como referência a produção industrial, que seguem na tabela:

[Tabela 9 aqui]

Os índices calculados tiveram resultados semelhantes aos de [Issler et al. \(2013\)](#), que obtiveram índices entre 20% e 30%. Ainda segundo os autores o índice é considerado como um bom preditor do estado da economia se $QPS < 0.4$. Para o índice de incerteza todos os valores estão abaixo de 0.4. Ressalta-se ainda que o índice para $h=12$ obteve a melhor performance, com um valor de 0.28.

A tabela 8 apresenta as datas de turning point para os índices de incerteza e para a produção industrial. Observa-se que em 2008 todos os índices de incerteza apresentam turning point por volta dos meses de abril e maio de 2008. Esse já seria um indício da recessão e de crise econômica que viria a ocorrer a partir de setembro de 2008 e em 2009. De forma análoga, ocorre também turning point em 2011 nos índices de incerteza, o que seria algum indício das políticas da nova matriz macroeconômica, indicando que a incerteza estava baixa, dado o crescimento ocorrido em 2010 e volta a aumentar a partir de 2011 devido a quebra do tripé macroeconômico. Outro ponto interessante é que o índice de incerteza $h = 3$ consegue captar o turning point de março de 2014, sinalizando um aumento da incerteza e um cenário econômico caracterizado pela presença de recessão e inflação elevada de 2014 em diante.

1.6 Considerações Finais

A pesquisa propôs a construção de um índice de incerteza macroeconômica para o Brasil baseado na metodologia proposta por [Jurado et al. \(2015\)](#). Para a construção do índice brasileiro foram utilizados as séries macroeconômicas do trabalho de [Issler et al. \(2013\)](#).

A medida de incerteza macroeconômica brasileira teve uma correlação negativa com a produção industrial, sendo esta correlação menor que a do IBOVESPA, indicando assim uma maior precisão na identificação antecipada de choques macroeconômicos negativos. De fato, a pesquisa encontrou uma relação de causalidade bem definida entre o índice de incerteza macroeconômica e variáveis macroeconômicas como a produção industrial. Foi verificado que incerteza causa produção industrial, mas produção industrial não causa incerteza. Desta forma, a incerteza precede a produção industrial no tempo, indicando que maiores níveis de incerteza causam uma queda da produção industrial no futuro.

Além da série de incerteza com pesos $1/N$, também foram construídos índices com pesos estimados com base na análise de componentes principais. Assim, foi dado um maior pesos as séries que supostamente tem maior influência sob a incerteza. Os índices com pesos baseado em componentes principais melhoram a performance do índice no que diz respeito a correlação, dado que os mesmos tem uma correlação menor com a produção industrial do que os índice obtido usando pesos $1/N$.

Também foi encontrado um alto índice de incerteza para os anos de 2014 e 2015, gerado principalmente pelo cenário econômico atual, onde os agentes econômicos esperam e veem como o futuro se desdobrar. Esta é a teoria de ciclos reais de negócios conhecida por “wait and see”. Nesse contexto os consumidores adiam decisões de consumo e empresários adiam decisões de investimento, tornando a queda do produto inevitável. De acordo com os resultados, esse movimento já fora captado pelo aumento do índice de incerteza a partir de 2014 e também pela relação de causalidade entre incerteza e produção industrial. Neste sentido, como sugestões de política econômica, recomenda-se o uso do índice de incerteza por parte de instituições do governo ou do mercado financeiro como um indicador antecedente de atividade econômica futura, como forma de captar uma possível recessão ou expansão econômica.

Para futuras pesquisas sugere-se também a construção de um índice baseado em pesquisas de mercado, como a pesquisa focus. O índice pode ser construído com base no erro de previsão ou surpresas a partir das previsões dos analistas de mercado, seguindo a metodologia de [Scotti \(2016\)](#). Seria interessante uma comparação da performance dos índices de incerteza para o mercado brasileiro construídos via erro de previsão gerado a

partir do modelo econométrico e via erro de previsão gerado a partir das expectativas dos analistas de mercado.

Capítulo 2

Identificando o Poder Preditivo das Minutas do FED

2.1 Introdução

De acordo com [Gentzkow et al. \(2017\)](#), a informação codificada no texto é um complemento rico para os tipos de dados mais estruturados tradicionalmente usados na pesquisa empírica. De fato, nos últimos anos, temos visto um intenso uso de dados textuais em diferentes áreas de pesquisa. A ideia consiste em transformar *strings variables* em variáveis numéricas e usá-las como preditores em diferentes modelos. Vários estudos já exploraram essa fonte adicional de informação. Na literatura financeira, [Garcia \(2013\)](#) estuda o efeito do sentimento sobre os preços dos ativos durante o século XX (1905 a 2005). O autor usa textos do New York Times para construir variáveis de sentimento usando análise textual. Ele conclui que a previsibilidade dos retornos das ações usando o conteúdo de notícias é particularmente forte durante as recessões. [Engelberg and Parsons \(2011\)](#) comparam o comportamento de investidores com acesso a diferentes coberturas da mídia do mesmo evento. Eles se concentram na relação entre a cobertura da mídia local e o volume de negociações da carteira de ações local e descobrem que a cobertura da mídia local é um forte preditor dos retornos das firmas locais.

Na literatura macroeconômica, [Dossani \(2018\)](#) analisa como o tom das conferências de imprensa do Banco Central dos Estados Unidos impacta os prêmios de risco no mercado de câmbio. Ele mede o tom como a diferença entre o número de frases “*hawkish*” e “*dovish*” feitas durante uma conferência de imprensa. Ele usou quatro contratos futuros de moeda negociados na bolsa de mercadorias e futuros de Chicago *Chicago Mercantile Exchange* (CME) e descobriu que a aversão ao risco implícita aumenta quando os

Bancos Centrais são agressivos (“hawkish”) e diminui quando os Bancos Centrais são “dovish”. Outros exemplos em que dados textuais são usados para análise macroeconômica incluem [Armesto et al. \(2009\)](#), [Boukus and Rosenberg \(2006\)](#), [Cecchetti et al. \(2003\)](#), entre outros.¹

Uma suposição comum entre os estudos acima é que o conteúdo do dicionário (conjunto de palavras) usado para construir índices de sentimento/informação é constante ao longo do tempo e subjetivamente escolhido pelo usuário. Na prática, isso implica que, dado um dicionário fixo, calculamos a frequência com que cada palavra (ou combinação de palavras) aparece nos textos e usamos essas informações para construir preditores (índices de sentimento/informação) que podem ser usados para previsão ou função de impulso resposta. A suposição de dicionário invariante no tempo é particularmente difícil de manter em documentos que introduzem novas palavras ao longo do tempo ou se o vocabulário usado em períodos de recessão difere daquele usado em períodos de expansões econômicas. Mesmo que o vocabulário fosse o mesmo ao longo do tempo, o poder preditivo de algumas palavras pode variar, mas a literatura existente não leva em conta esse efeito e, portanto, os preditores resultantes não refletem as informações textuais mais preditivas encontradas nos documentos em um determinado momento.

Neste artigo, permitiremos que o conteúdo dos dicionários varie ao longo do tempo, tornando-o inteiramente determinado pelo poder preditivo de suas palavras. O objetivo é, portanto, maximizar o poder preditivo do dicionário. Assim propomos novos preditores que dependem de dicionários que variam com o tempo e mostramos que eles têm mais poder preditivo do que aqueles construídos a partir de dicionários invariantes no tempo. Nossa metodologia pode ser resumida em três etapas: (i) Na primeira etapa, um vetor X_t é criado, onde cada elemento mostra a frequência em que cada palavra (ou combinação de palavras) aparece na minuta do FED no tempo t . Assim, este passo transforma palavras em valores numéricos sem usar um dicionário pré-especificado; (ii) na segunda etapa, usamos o aprendizado de máquina (“*Machine Learning*”) (“*Elastic Net*”) para identificar um subconjunto $X_t^* \subset X_t$. Esta é a representação numérica do nosso dicionário mais preditivo que é então usado para construir um preditor (ou conjunto de preditores), $D_t = g(X_t^*)$. Por exemplo, a função $g(\cdot)$ pode selecionar apenas as palavras com maior poder preditivo k^{th} ; também pode representar fatores comuns (calculados por componentes principais) de palavras em X_t^* ; ou podemos até deixar $g(\cdot)$ representar um índice de sentimento baseado no dicionário mais previsível X_t^* . Na seção 3, várias possibilidades para $g(\cdot)$ são consideradas; (iii) finalmente, na terceira etapa, usamos o(s) preditor(es) D_t para prever nossa variável de interesse. Este procedimento de 3 passos

¹[Wright \(2012\)](#) e [Altavilla and Giannone \(2017\)](#) estudam o efeito das notícias sobre a política monetária na curva de juros, mas não dependem da mineração de textos (*text mining*). Eles encontraram que as notícias afetam as expectativas dos agentes de mercado sobre os rendimentos das obrigações das empresas e dos títulos do Tesouro.

é repetido recursivamente até final da amostra (previsão recursiva), implicando que o conteúdo do dicionário com maior poder de previsão muda ao longo do tempo².

Nosso artigo está relacionado, em motivação, ao trabalho recente de [Thorsrud \(2018\)](#) e [Hansen et al. \(2017\)](#) que usaram a Latent Dirichlet Allocation(LDA) para endereçar limitações impostas por dicionários pré-especificados. A principal diferença entre nossa abordagem e a baseada no LDA é que a última é uma técnica de aprendizado de máquina não supervisionado, ao passo que nossa abordagem se baseia no aprendizado de máquina supervisionado. Como explicado em [Chakraborty and Joseph \(2017\)](#), a diferença entre as duas abordagens é a existência de uma variável dependente y . O aprendizado supervisionado é o caso clássico de modelagem y usando entradas (preditores) x em que escolhemos um algoritmo de aprendizado que se ajustará ao objetivo usando os recursos de entrada fornecidos. No aprendizado não supervisionado, não existe um alvo y e os algoritmos apenas visam encontrar estrutura nos dados, por exemplo, agrupando observações ou agrupando palavras em tópicos, como feito pela técnica LDA. Estimação recursiva é um componente chave para gerar dicionários variantes no tempo pelo método proposto neste artigo. De fato, nossa metodologia consiste em escolher recursivamente um conjunto de palavras que minimiza uma função perda estritamente convexa, implicando que apenas as palavras mais preditivas serão incluídas no dicionário a cada momento t . Como mostrado por [Thorsrud \(2018\)](#), a abordagem do LDA é muito intensiva em computação e a atualização recursiva da amostra de treinamento para re-estimar os tópicos do modelo é inviável.³ Nossa metodologia, por outro lado, permitirá uma atualização recursiva de baixo custo da amostra de treinamento para re-estimar o modelo de previsão usando técnicas de SML.⁴

A metodologia proposta tem várias inovações. Em primeiro lugar, as palavras (representação numérica) contidas em X_t^* , em vez de serem escolhidas pelo usuário, são selecionadas com base em seu poder preditivo. Isto é realizado através da aplicação de Aprendizado de Máquina Supervisionado no segundo passo para selecionar as palavras com maior poder preditivo. Segundo, ao contrário da metodologia existente, onde o conteúdo do dicionário é fixo ao longo do tempo, implicando que a variação no tempo do preditor D_t é explicada unicamente pela frequência com que cada palavra pré-especificada aparece em um documento, a nova metodologia também permite variação no tempo no conteúdo do dicionário. Mostramos que essa inovação adicional tem potencial para melhorar substancialmente as previsões macroeconômicas fora da amostra. No nosso artigo,

²No Aprendizado de Máquina Supervisionado a validação do modelo é essencial para garantir a performance fora da amostra do modelo. Discutimos na seção 2.3.2 a validação do modelo em detalhes.

³Essa restrição levou [Thorsrud \(2018\)](#) a manter constante a amostra de treinamento.

⁴Outra ressalva com a reestimativa do LDA recursivamente é a falta de identificabilidade, ou seja, estimativas dos tópicos não podem ser combinadas entre amostras para uma análise que dependa do conteúdo de tópicos específicos ([Thorsrud, 2018](#), p. 22).

usamos preditores baseados na minutas do FED para prever o Produto em tempo real.⁵ Os bancos centrais contam com previsões de produção para tomar decisões sobre mudanças nas políticas monetárias, enquanto o setor privado usa as previsões de crescimento do Produto para tomar decisões sobre investimento, marketing e gerenciamento de risco. Um grande conjunto de modelos de previsão de crescimento do Produto foram sugeridos recentemente pela literatura, que inclui modelos autorregressivos lineares simples (AR) bem como previsões dos analistas e métodos mais sofisticados baseados no Modelo Dinâmico de Equilíbrio Geral Estocástico⁶.

Para ser claro, a principal contribuição metodológica do artigo não é propor uma previsão final para o Produto mas, ao invés disso combinar novos desenvolvimentos a partir da regressão de texto e Aprendizado de Máquina Supervisionado um uma maneira inédita para identificar a informação mais preditiva das minutas do FED. Para ilustrar o poder preditivo das minutas do FED, escolhemos um $AR(2)$ linear simples para representar o modelo de previsão benchmark. Nossa escolha é motivada pelos resultados de [Chauvet and Potter \(2013\)](#) que compararam o desempenho de muitos modelos econométricos usados para prever o crescimento do PIB dos EUA de 1992 a 2010 e concluíram que um modelo $AR(2)$ simples tem o melhor desempenho durante os períodos de expansão e funciona relativamente bem durante as recessões. Suas descobertas sustentam evidências anteriores de que as previsões de crescimento de produção a partir de modelos autorregressivos simples são difíceis de superar. Por esse motivo, escolhemos um modelo simples $AR(2)$ como o benchmark mais forte a ser superado pelo método proposto. Nossos resultados indicam que os dados textuais encontrados em minutas do FED contêm poder preditivo não-trivial, mas os dicionários fixos não são a melhor abordagem para identificar tal informação. De fato, quando adicionamos preditores baseados em dicionários fixos ao benchmark $AR(2)$, não encontramos nenhuma melhora nas previsões, sugerindo que os preditores computados de um dicionário fixo não adicionam poder preditivo ao modelo de benchmark. No entanto, quando esses preditores são calculados a partir de dicionários que variam com o tempo, encontramos um forte ganho de precisão em relação ao modelo aninhado(“nested”) $AR(2)$. Também reportamos que as previsões baseadas nas minutas do FED propostas performam melhor do que as previsões do green book do FED⁷, sugerindo que a informação gerada durante a reunião do FOMC contém um poder de previsão não trivial.

⁵O uso de dados em tempo real ao invés da revisões do PIB implica que estamos considerando apenas a informação que estava disponível naquele momento que a previsão estava sendo feita. Assim, estamos reproduzindo o problema de previsão em tempo real.

⁶Dynamic Stochastic General Equilibrium Model(DSGE)

⁷O green book ou livro verde do FED é o livro onde estão as previsões do staff do FED sobre as principais variáveis macroeconômicas da economia. Os membros do FED baseiam-se nas previsões do green book para tomar sua decisão de política.

A escolha do Aprendizado de Máquina Supervisionado(shrinkage)⁸ também importa. Mostramos que dicionários variantes no tempo que dependem do elastic net para selecionar as palavras mais preditivas tem maior poder preditivo do que aqueles que dependem do LASSO ou ridge regression. Uma explicação para este resultado é que ridge regression não executa a seleção do modelo(não encolhe os coeficientes para zero) e, diferentemente do LASSO, o número de preditores(palavras) selecionados pelo elastic net não é limitado pela tamanho da amostra T . Além do mais, elastic net é também robusto a efeitos de grupo, que é particularmente importante para nossa análise devido muitas palavras que aparecem nas minutas do FED, tais como “inflation” e “unemployment” serem altamente correlacionadas.⁹ Assim, elastic net não apenas seleciona as palavras mais preditivas, mas também garante que palavras igualmente preditivas não serão descartadas aleatoriamente apenas porque são altamente correlacionadas com outras palavras preditivas. Na literatura de previsão, um curta lista de artigos que com sucesso empregaram elastic net incluem [Bai and Ng \(2008\)](#), [Li et al. \(2015\)](#) e [Lima et al. \(2018a\)](#). Nossos resultados sugerem que a previsão macroeconômica poderia ser melhorada através da combinação de dados textuais e aprendizado de máquina de forma que apenas as palavras mais preditivas fossem consideradas para a análise de previsão e na função impulso resposta. Isso tem uma implicação significativa para aqueles interessados em avaliar os efeitos da política monetária na economia. A identificação de choques na política monetária continua a ser um desafio importante devido a evidências inconclusivas relatadas pelos principais pesquisadores. Veja [Christiano et al. \(1999\)](#) e [Ramey \(2016\)](#) para uma pesquisa detalhada. Os preditores baseados em texto identificados neste estudo poderiam potencialmente capturar verdadeiros choques de política monetária para avaliar os efeitos da política monetária sobre o produto.

O restante deste artigo está organizado da seguinte forma. A seção 2.2 introduz o canal através do qual as informações das atas do FED afetam os valores futuros do crescimento do produto. A seção 2.3 apresenta a regressão de texto (*Text Regression*) e explica como usamos o aprendizado de máquina para identificar um dicionário que varia no tempo. A seção 2.4 explica como definiremos os preditores que usaremos em nosso exercício de previsão. A seção 2.5 introduz nossa análise empírica, incluindo uma descrição completa do conjunto de dados, os modelos de previsão e os métodos usados para avaliar as previsões fora da amostra. Seção 2.6 apresenta os principais resultados seguidos de considerações finais na seção 2.7.

⁸No nosso caso Elastic Net, LASSO e Ridge Regression.

⁹palavras são positivamente(negativamente) correlacionadas quando o número de vezes que elas aparecem em um documento é positivamente(negativamente) correlacionado ao longo do tempo.

2.2 Decisões do Banco Central, curva de juros e crescimento do produto

Nesta seção, explicamos o canal através do qual as informações contidas nas atas do FED afetam o crescimento econômico. A identificação de indicadores úteis para prever atividades econômicas futuras é uma busca importante para formuladores de políticas e vários investidores. A principal justificativa explorada neste estudo é que o tom do banco central em relação às políticas monetárias futuras atuais e esperadas (como demonstrado nas reuniões e anúncios do FED) tem poder explicativo significativo na previsão do crescimento futuro do produto.

Antes de destacar o mecanismo de transmissão subjacente, é preciso recordar os principais objetivos do banco central. Por meio de ações de política monetária, o FED deseja alcançar crescimento econômico sustentável e estabilidade de preços. Para cumprir com sucesso seu mandato duplo, o FED se envolve em políticas monetárias contracíclicas. Por exemplo, para combater contrações econômicas ou contrações econômicas esperadas, ela adota uma política monetária expansionista. Da mesma forma, durante o boom econômico, o FED provavelmente buscará políticas contracionistas para controlar a estabilidade de preços. Isto é particularmente verdadeiro quando a taxa de inflação atual e esperada estão acima da meta de inflação do FED. A eficácia de várias ações de política monetária depende, em última instância, da capacidade do FED de controlar com êxito as expectativas do setor privado sobre taxas de juros e inflação futuras. A taxa de juros real esperada de longo prazo é a variável mais relevante para as empresas em suas decisões de investimento. Isto também é verdade para os agregados familiares nas suas decisões de gastos em consumo durável. Assim, para o FED controlar a taxa de juros real de longo prazo, ele deve ser capaz de controlar a taxa de juros nominal de longo prazo ou as expectativas de longo prazo da taxa de inflação, ou ambas. Por meio de anúncios de política monetária em termos de intensidade e direções (tom), o banco central é capaz de controlar tanto as taxas de juros de curto prazo quanto de longo prazo (*yield spread e term structure*) que afeta, em última instância a atividade econômica.

Muitas pesquisas empíricas no passado identificaram a inclinação do *yield spread* como um bom preditor do crescimento do produto. O *yield spread* mede a diferença entre as taxas de juros de curto e longo prazo. Sob condições de mercados eficientes, espera-se que o gap do spread contenha as expectativas do setor privado em relação às condições econômicas futuras, bem como às futuras políticas governamentais. Normalmente, os títulos de longo prazo têm rendimentos mais elevados do que os títulos de curto prazo, indicando uma curva de juros inclinada para cima. Da mesma forma, uma curva de rendimentos invertida ou plana indica que as atuais taxas de juros de curto prazo são mais

altas do que as taxas de longo prazo. Uma curva de juros positivamente inclinada com spread maior tende a prever um crescimento mais rápido do produto no futuro. Uma curva de rendimentos invertida ou plana, por outro lado, levanta preocupações porque indica uma desaceleração econômica e sinaliza uma recessão econômica. Estudos importantes que testaram este canal para prever o crescimento econômico (principalmente recessões econômicas) incluem [Ahrens \(2002\)](#), [Ang et al. \(2006\)](#), [Bernard and Gerlach \(1998\)](#), [Duarte et al. \(2005\)](#), [Estrella \(2005\)](#), [Estrella and Hardouvelis \(1991\)](#), [Estrella and Mishkin \(1998\)](#), [Plosser and Rouwenhorst \(1994\)](#), [Stock and Watson \(1989\)](#) entre outros.

É claro que o banco central conduz a política monetária para afetar as taxas de juros nominais de curto prazo, a taxa de inflação e as taxas de juros reais de longo prazo para alcançar seus objetivos. Os efeitos dos choques da política monetária convencional (ou seja, a taxa de crescimento da oferta monetária ou da meta de inflação) na estrutura a termo das taxas de juros são bem conhecidos. Veja [Mansoorian and Mohsin \(2004\)](#) para detalhes. Neste estudo, estamos interessados em avaliar os efeitos da tonalidade da autoridade monetária. Frequentemente, o FED comunica suas mudanças futuras de políticas desejadas por meio de reuniões do FOMC e conferências de imprensa. Desde o início da década de 2000, o banco central começou a usar essa orientação em relação ao futuro como uma ferramenta de política efetiva. Através desta ferramenta política, o FED espera se comunicar com o setor privado em relação as suas ações futuras. Com forte credibilidade, o banco central poderia efetivamente e com sucesso influenciar a expectativa do setor privado sobre o futuro ([Kydland and Prescott, 1977](#)) e, portanto, afetar os valores futuros do crescimento do produto. Após a grande recessão de 2007-2008, quando as taxas dos fundos federais (*Federal Fund Rates*) permaneceram próximas de zero, a comunicação do FED através de seus relatórios mensais de política mostraram-se muito eficazes no gerenciamento das expectativas. Efetivamente, a comunicação do FED controla a estrutura a termo das taxas de juros reais na economia. A análise impulso resposta fornecida por [Lucca and Moench \(2015\)](#) e por nós mesmos no final deste artigo demonstrou que a tonalidade do FED afeta de fato significativamente os spreads.

A discussão acima sugere que o tom do FED pode prever o crescimento do produto por meio de mudanças no yield spread. Neste artigo, propomos usar uma abordagem de dicionário que varie no tempo para identificar o tom do banco central selecionando corretamente as palavras mais preditivas das minutas do FED. Construiremos novos preditores a partir do dicionário que varia no tempo e os usaremos para prever o crescimento da produção dos EUA. A ideia principal dessa nova abordagem é que é possível maximizar o poder preditivo do dicionário, permitindo que seu conteúdo mude com o tempo. Na próxima seção, explicamos brevemente como a regressão de texto e o aprendizado de máquina podem ser combinados para construir dicionários que variam no tempo.

2.3 Text Regression

A regressão de texto é um ramo da econometria que consiste em transformar as variáveis de texto em numéricas. Conforme documentado em [Gentzkow et al. \(2017\)](#), tem havido um interesse crescente nas informações codificadas no texto como um complemento rico para os tipos de dados mais estruturados tradicionalmente usados na pesquisa acadêmica. Nesta seção, explicaremos como os dados textuais entram na nossa base de dados e como métodos de regularização como o “Elastic Net” podem ser usados para computar um dicionário variante no tempo.

2.3.1 Dados Textuais

Nesta seção explicamos como os dados textuais entram na base de dados. Nosso objetivo inicial é transformar o texto em dados numéricos. Para realizar essa tarefa, primeiro salvamos todos os textos baixados no que a literatura chama de “um corpus”, que é uma coleção de textos escritos, ou seja, um conjunto de textos. Antes de realizar a contagem de palavras, realizamos o pré processamento do texto bruto em vários passos. O objetivo é reduzir o vocabulário a um conjunto de palavras que são mais significativas. Seguindo [Hansen et al. \(2017\)](#), primeiro identificamos colocações ou sequências de palavras que possuem um significado específico. Por exemplo, “federal fund”, “financial market”, “labor market” correspondem a um conceito econômico único, mas são compostos de duas palavras separadas. Assim, identificamos as colocações usando os taggers propostos por [Toutanova et al. \(2003\)](#) e criamos um termo único para sequências de duas palavras cuja frequência é superior a 100. Também removemos pontuação, stopwords, etc. Esse processo de limpeza foi realizado através dos seguintes comandos:

tolower: Como o software R faz distinção entre letras maiúsculas e minúsculas, usamos o comando “tolower” para classificar palavras com letras maiúsculas e minúsculas (como house e House) como palavras iguais.

removePunctuation: este comando foi usado para remover a pontuação.

removenumbers: Usamos este comando para remover números.

stopwords: Usamos este comando para remover palavras como stop, that, which, what, etc.

stripWhitespace: Este comando é usado para remover múltiplos espaços, de modo que “economia” e “economia ” sejam tratados como tendo o mesmo significado.

stemming: Esta função é usada para garantir que a DTM(document term matrix) inclua apenas os radicais da palavra. Por exemplo, palavras como Economy e Economics são contadas como Econom. Se ambos aparecerem em um documento, sua soma é mostrada. Caso contrário, uma única palavra é contada. Neste artigo, consideramos apenas os radicais das palavras.

Em nosso estudo de previsão de crescimento do Produto, dividimos a amostra total $T = R + P$ em porções dentro da amostra e fora da amostra. As observações dentro da amostra vão de 1 a R , enquanto que as previsões fora da amostra vão de $R + 1$ até T , e temos um total de P previsões h -passos a frente fora da amostra. Haverá um corpus para cada trimestre s , $s = 1, \dots, t - h$ com $t = R, \dots, T - h$ em que h é o horizonte de previsão.¹⁰ Como a contagem ocorre em todas as minutas disponíveis no trimestre s , terminamos com a nova série temporal $X_{j,t} = \{X_{j,s}\}_{s=1}^t$ para cada palavra $j = 1, \dots, p$ e ponto de partida das previsões t . Assim, nossa contagem de palavras gera novas séries temporais que contêm observações até o momento (ponto de partida das previsões) t , com $t = R, \dots, T - h$.

A fim de evitar problemas relacionados à estimação de fatores comuns por componentes principais, normalizamos a nova série temporal usando sua respectiva média histórica e desvio padrão. Especificamente, calculamos $(X_{j,t} - \mu_j) / \sigma_j$ onde μ_j e σ_j são a média histórica e desvio padrão, respectivamente, da palavra j computada usando observações até o momento t . Assim, antes de fazer uma nova previsão, contamos as palavras(ou combinação de palavras) sem impor qualquer dicionário pré-especificado e normalizamos a série temporal resultante.¹¹ Nossa etapa final é coletar todas as séries temporais normalizadas em vetor $p \times 1$, $X_t = (X_{1,t}, \dots, X_{p,t})'$ onde a dimensão do vetor X_t é igual ao número de palavras que aparecem nas minutas, ou seja, $\dim(X_t) = p$. A figura (13) exibe uma série temporal correspondente a uma contagem normalizada da colocação (termo de duas palavras) “economic activity”. A série temporal varia de 1958.T1 a 2017.T2 e as barras cinzas indicam períodos de recessão econômica. É interessante notar que o termo “economic activity” foi amplamente empregado durante as reuniões do FOMC que ocorreram em recessões (especialmente em 1975 e 2008). Isso sugere que o termo pode ser especialmente útil para prever o crescimento do produto durante os períodos de recessão. Em nossa seção empírica, mostramos que esse termo é de fato selecionado como um importante preditor do crescimento do produto¹² durante os períodos de recessão.

¹⁰Um corpus em trimestre s incluirá todas as minutas do FED daquele trimestre.

¹¹A normalização implica que se uma palavra não aparecer na ata durante o trimestre s , então ele receberá um valor $(-\mu_j) / \sigma_j$. Observe, no entanto, que o pré-processamento de textos brutos, explicado nesta seção, remove de nossa amostra palavras que são muito raras, através do cálculo do TF-IDF. Isso evita a ocorrência de observações que são quase sempre iguais a $(-\mu_j) / \sigma_j$.

¹²Nossa variável dependente, a taxa de crescimento do produto anualizado, é definida como $y_t = \ln(\frac{Y_t}{Y_{t-1}}) * 400$, onde Y_t representa o Produto Interno Bruto Real da economia dos Estados Unidos.

{Coloque a figura 13 aqui. }

2.3.2 Construindo Dicionários que variam no tempo

Nesta seção, queremos construir um dicionário que varie no tempo, isto é, uma matriz $X_t^* \subset X_t$ que contém apenas as palavras com maior poder preditivo das minutas do FED. Nossa variável dependente é a taxa de crescimento do produto dos Estados Unidos em tempo real y_t . Implementamos este passo aplicando *Elastic Net* na seguinte equação de predição linear:

$$y_t = W'_{t-h}\beta_h + X'_{t-h}\phi_h + \epsilon_t \quad (2.1)$$

onde $h \geq 0$ é o horizonte de previsão, $\hat{\beta}_h$ e $\hat{\phi}_h$ são estimados pela seguinte função objetivo:

$$\min_{\beta_h, \phi_h} \sum_t (y_t - W'_{t-h}\beta_h - X'_{t-h}\phi_h)^2 + \lambda_1 \|\phi_h\|_{\ell_1} + \lambda_2 \|\phi_h\|_{\ell_2} \quad (2.2)$$

onde W_t é um vetor $k \times 1$ de preditores predeterminados, como lags de y_t , bem como preditores tradicionais de dados estruturados; X_t é um vetor $p \times 1$ definido anteriormente, e $\|\cdot\|_{\ell_1}$ e $\|\cdot\|_{\ell_2}$ são as normas ℓ_1 e ℓ_2 , respectivamente. A norma $\|\cdot\|_{\ell_1} = \sum_{j=1}^p |\phi_h|$ penaliza o valor absoluto dos coeficientes ao passo que $\|\cdot\|_{\ell_2} = \sum_{j=1}^p \phi_h^2$ penaliza os coeficientes ao quadrado. A equação 2.2 é estimada recursivamente até o fim da amostra, ou seja, regredimos as observações y_{s+h} sobre os preditores W_s e X_s para $s = 1, \dots, t-h$ e $t = R, \dots, T-h$.¹³

Na literatura de estatística, uma combinação da norma restrição ℓ_1 e da norma de restrição ℓ_2 é conhecido como “Elastic Net” (Zou and Hastie, 2005). Como não estamos considerando um dicionário fixo, a dimensão de X_t será muito grande, isto é, $p \gg T$, e por essa razão ϕ_h serão os únicos coeficientes penalizados em (2.2). A escolha das palavras mais preditivas depende dos valores dos parâmetros de ajuste λ_1 e λ_2 . Nesta tese nós estimamos os parâmetros de ajuste usando procedimento de validação cruzada (cross validation) sugerido no pacote GLMNET R (seção de regressão linear) desenvolvido por Trevor Hastie e Junyang Qian.¹⁴ Este procedimento de validação cruzada executa uma seleção ex ante dos parâmetros de ajuste que é essencial para evitar overfitting. Em particular, empregamos a validação cruzada para dados dependentes como em Elliott

¹³Paramos em $T-h$ porque precisamos da observação T para avaliar a previsão.

¹⁴O link pode ser encontrado aqui [GLMNET](#).

and Timmermann (2013a) e Diebold and Shin (2018). Quando $\lambda_1 = \lambda_2 = 0$ então a função objetivo (2.2) torna-se igual à soma usual dos resíduos quadrados. Quando $\lambda_2 = 0$, impomos apenas a norma de restrição ℓ_1 e temos o estimador LASSO, ao passo que quando $\lambda_2 = 0$ temos ridge regression. Ridge regression não faz a seleção de modelos, dado que não reduz os coeficientes em torno do zero. Uma das vantagens do Elastic Net em relação ao LASSO é que o mesmo pode selecionar um número de variáveis maior que o tamanho da amostra t .

Elastic Net corresponde ao caso em quem ambos λ_1 e λ_2 são positivos. Nesse caso, $\hat{\phi}_h$ é reduzido para zero de duas maneiras diferentes, promovendo tanto a dispersão quanto a estabilidade. Em outras palavras, o Elastic Net é um estimador que reduz todas as estimativas de parâmetros em torno do zero, enquanto penaliza mais os parâmetros menores e maiores. É importante notar que isto é conseguido de uma forma que reduz o erro quadrado médio (MSE) do estimador de parâmetros e dos erros de previsão do modelo, sugerindo que podemos usar Elastic Net para selecionar um vetor com a série temporal mais preditiva (palavras) X_t^* . Especificamente, as séries temporais incluídas em X_t^* são aquelas cujos coeficientes estimados ϕ_h na equação (2.2) são diferentes de zero. Desde que a estimação ocorri recursivamente sobre $t = R, \dots, T - h$, diferentes séries de tempo (palavras) $X_t^* \subset X_t$ serão selecionadas ao longo do tempo. Isto é o que nós chamamos dicionário variante no tempo. É importante observar que o conteúdo do dicionário variante no tempo também depende do horizonte de previsão h , mas omitimos esta notação extra para facilitar a exposição. Na prática, esta abordagem permite ao econometrista ter uma vantagem de usar o dicionário mais preditivo (X_t^*) a cada ponto de partida da previsão t para efetuar a previsão h -passos a frente.

Embora LASSO seja bem sucedido na seleção de variáveis, no caso particular em que o número de preditores é maior que o tamanho da amostra T , como no nosso exercício empírico, o LASSO seleciona no máximo T preditores antes de saturar. Por esta razão Zou and Hastie (2005) descreveram a combinação da restrição- ℓ_1 e restrição- ℓ_2 como “uma rede de pesca extensível que retém todo o peixe grande”. Além do mais, em um problema de regressão se a i -ésima coluna e a j -ésima coluna de X_t são altamente correlacionadas e ambos são variáveis (palavras) independentes importantes, a regressão com apenas a norma de restrição- ℓ_1 tende a atribuir uma grande estimativa para uns dos $\phi_{i,h}$ e $\phi_{j,h}$ e aleatoriamente, atribui ao outro o valor de zero (Zou and Hastie, 2005, Efron et al., 2004). Mas com a norma de restrição adicional ℓ_2 , a regressão com ambas as restrições tende a produzir estimativas similares para $\phi_{i,h}$ e $\phi_{j,h}$ enquanto mantém a dispersão. Isto é particularmente importante para a nossa análise porque muitas palavras que aparecem nas minutas do FED são altamente correlacionadas.¹⁵ Assim, elastic net não apenas seleciona as palavras mais preditivas, mas também garante que palavras igualmente

¹⁵veja nota de rodapé 9.

preditivas não serão aleatoriamente descartadas porque elas são correlacionadas com outras palavras preditivas.

O próximo passo é construir o preditor baseado na minuta do FED $D_t = g(X_t^*)$ onde X_t^* foi selecionado como demonstrado acima. Nós finalmente adicionamos o preditor baseado no FED para a equação de predição linear e o usamos para efetuar a previsão recursiva h-passos a frente. Em outras palavras, consideramos a seguinte equação de previsão.

$$y_{t+h} = W_t' \beta_h + D_t' \phi_h + \epsilon_{t+h} \quad (2.3)$$

e computamos as previsões h-passos a frente como:

$$f_{t+h,t} = W_t' \hat{\beta}_h + D_t' \hat{\phi}_h \quad (2.4)$$

onde $f_{t+h,t}$ é a previsão h-passos a frente de y_{t+h} feita no tempo $t = R, \dots, T-h$; $\hat{\beta}_h$ e $\hat{\phi}_h$ são as estimativas feitas por MQO. Recorde que dividimos a amostra total de $T = R + P$ observações em porções dentro da amostra e fora da amostra. As observações dentro da amostra vão de 1 até R , ao passo que as observações fora da amostra vão de $R + 1$ até T , para um total de P h-passos a frente previsões fora da amostra. Então, para cada ponto de partida da previsão $t = R, \dots, T-h$, estimamos os coeficientes $\hat{\beta}_h$ e $\hat{\phi}_h$ regredindo as observações y_{s+h} sobre os preditores W_s e D_s para $s = 1, \dots, t-h$ (esquema de previsão recursiva), e computamos $f_{t+h,t}$ pela equação de avaliação (2.3) na t -ésima observação. Em seguida, descrevemos os preditores baseados no FED usados na seção empírica do artigo.

2.4 Preditores baseados nas Minutas do FED

2.4.1 Dicionários variantes no tempo

Nesta seção, usaremos as séries temporais selecionadas (palavras) X_t^* para construir preditores de variáveis macroeconômicas, como a taxa de crescimento do Produto. Todos os preditores podem ser expressos como $D_t = g(X_t^*)$ e, portanto, podemos considerar várias opções. A seguir, classificamos um modelo de previsão baseado no tipo de preditor utilizado.

$(D_{1,t})$: Nosso primeiro conjunto de preditores correspondem a fatores comuns de X_t^* , que são estimados por componentes principais usando os dados $X_{k,s}^*$ até a origem da

previsão t , onde $k = 1, \dots, K$ são as séries de tempo com maior poder preditivo $K \ll p$.¹⁶ Em outras palavras, os estimadores de componentes principais são definidos como:

$$(\Lambda_t, F_t) = \arg \min_{\{\lambda_k, f_s\}} \frac{1}{Kt} \sum_{k=1}^K \sum_{s=1}^t (X_{k,t}^* - \lambda_k' f_s)^2 \quad (2.5)$$

onde $\Lambda_t = (\lambda_{1,t}, \dots, \lambda_{K,t})'$ é a $K \times r$ matriz de loading e os r fatores comuns correspondentes são coletados na matriz $F_t = (f_{1,t}, \dots, f_{s,t})'$ de dimensão $t \times r$ onde $f_{s,t}$ denota a sth observação no vetor de fatores comuns estimados $r \times 1$ usando dados até o tempo t . O índice duplo (s, t) é com o propósito de explicitar que a informação que o r fatores comuns são estimados usando a informação até a origem da previsão t . Desta maneira, podemos usar as observações sobre os r fatores comuns disponíveis até a origem da previsão t para efetuar a previsão h -passos a frente. Esta abordagem é o padrão na literatura de previsão (veja, por exemplo (Gonçalves et al., 2017)).

Selecionamos o número ótimo de fatores por meio da abordagem de taxa de autovalores desenvolvida por Ahn and Horenstein (2013). Em seguida, seguimos Bai and Ng (2008) para manter apenas os fatores com p valor menor ou igual a 0.01. Por exemplo, se o número ideal é 3, então consideramos os três primeiros componentes principais, mas mantemos apenas aqueles com valor de p menor ou igual a 0,01. Essa abordagem é próxima da desenvolvida por Bai and Ng (2008), que aplicou a mesma ideia a dados estruturados, ou seja, propuseram primeiro selecionar as variáveis mais preditivas de um grande conjunto de covariáveis e, então, consideraram os fatores comuns das variáveis selecionadas como preditores. Observe que essa abordagem não impõe nenhuma restrição à porcentagem de informações “hawkish” e “dovish” carregadas em cada fator. De fato, mesmo que a quantidade de informações “hawkish” (“dovish”) nas minutas do FED não se altere com o tempo, essa abordagem é capaz de selecionar diferentes porcentagens de tais informações a serem carregadas nos fatores baseados em palavras selecionadas por Elastic Net. Se a informação hawkish tiver mais poder de previsão em um determinado momento, as palavras relacionadas ao sentimento hawkish serão selecionadas e os fatores serão carregados com mais informações hawkish. Neste artigo, um modelo de previsão que emprega esse tipo de preditores será rotulado como M_1 .

$(D_{2,t})$: Este conjunto de preditores é calculado a partir de bigramas (tokens) selecionados em vez de palavras. Bigramas são combinações de duas palavras que são usadas para evitar a ambiguidade semântica de palavras isoladas quando são retiradas do contexto. Seguindo Apel and Grimaldi (2012) consideramos uma lista de substantivos: (inflation, cyclical position, growth, price, wages, oil price, development, output

¹⁶Observe que $X_{k,s}^*$ é o k -ésimo elemento do vetor X_s^* .

growth, labor market); uma lista de adjetivos dovish: (decreasing, decreased, slower, weaker, lower, weak, low, slow); e adjetivos hawkish: (increasing, increased faster, stronger, higher, fast, strong, high).

Nós combinamos os substantivos e adjetivos listados acima, o que resulta em 192 bigramas (96 hawkish e 96 dovish). Em seguida, coletamos a frequência com que esses 112 bigramas aparecem na minuta do FED no momento t , dando origem a um vetor, B_t , com 112 novas séries temporais. Substituímos X_t por B_t em equações (2.1) e usamos Elastic Net (equação 2.2) para selecionar os bigramas mais preditivos B_t^* . Finalmente, calculamos os fatores comuns de B_t^* usando componentes principais. Selecionamos o número ideal de fatores por meio da abordagem *enginevalue ratio* desenvolvida por [Ahn and Horenstein \(2013\)](#) e mantemos apenas os fatores com valor de p menor ou igual a 0,01. Esses fatores selecionados serão os preditores usados para prever variáveis macroeconômicas. Neste artigo, um modelo de previsão que emprega esse conjunto de preditores será rotulado como M_2 . Observe que a principal diferença entre $D_{1,t}$ e $D_{2,t}$ é que o primeiro se baseia em palavras selecionadas, enquanto o segundo usa bigramas selecionados.

($D_{3,t}$): Para calcular esse conjunto de preditores, dividimos o mesmo conjunto de bigramas B_t usado em $D_{2,t}$ em um subconjunto positivo (dovish) (substantivos + adjetivos dovish) e um negativo (hawkish) (substantivos + adjetivos hawkish). Em seguida, aplicamos Elastic Net em ambos para selecionar os bigramas positivos mais preditivos, $B_t^{dovish*}$ e os bigramas negativos com mais previsão, $B_t^{hawkish*}$. Finalmente, aplicamos os componentes principais para obter um fator comum para $B_t^{dovish*}$ e um fator comum para $B_t^{hawkish*}$. Estes dois fatores comuns únicos serão nossos dois preditores. Observe que, ao contrário de M_1 e M_2 , essa abordagem impõe uma restrição na porcentagem de informações “hawkish” e “dovish” carregadas nos fatores. De fato, os fatores calculados usando os bigramas positivos são inteiramente carregados com informações “dovish”, enquanto os fatores calculados usando os bigramas negativos são carregados apenas com informações “hawkish”. Neste artigo, um modelo de previsão que emprega esse conjunto de preditores será rotulado como M_3 .

2.4.2 Dicionários fixos

($D_{4,t}$): Diferentemente dos preditores definidos anteriormente, este preditor é baseado em um dicionário fixo proposto por [Lucca and Trebbi \(2009\)](#). Eles consideram uma lista de adjetivos “hawkish” e “dovish”. **Lista de palavras hawkish** = {hawkish, tighten, hike, raise, increase, boost}; **Lista de palavras dovish** = {dovish, ease, cut, lower, decrease, loose}.

Procedemos fazendo a contagem das palavras e normalizando os dados. Então agregamos seguindo [Apel and Grimaldi \(2012\)](#) para calcular o índice:

$$D_{4,t} = \left[\left(\frac{hawk}{hawk + dove} \right) - \left(\frac{dove}{hawk + dove} \right) + 1 \right] \quad (2.6)$$

onde hawk e dove são o número de termos hawkish e dovish, e 1 é adicionado para excluir números negativos. Esse preditor pode ser interpretado como a hawkishness líquida da minuta no trimestre t . Nesta tese, um modelo de previsão que emprega esse preditor será rotulado como M_4 .

($D_{5,t}$): Os preditores desenvolvidos aqui também dependem de um dicionário fixo, mas usamos Elastic Net para selecionar as palavras mais preditivas do dicionário fixo. Deste modo, este método pode ser interpretado como um fusão entre dicionários variáveis e fixos no tempo. Em outras palavras, consideramos a mesma lista de palavras de $D_{4,t}$, então fazemos a contagem de palavras e normalizamos a série temporal resultante, levando-nos a dois conjuntos de séries temporais: o conjunto de hawkish $X_t^{hawkish}$ e X_t^{dovish} . O passo final é usar Elastic Net para identificar as palavras mais preditivas, isto é, um subconjunto de $X_t^{*hawkish}$ e um subconjunto de $X_t^{*dovish}$, respectivamente. Esses dois subconjuntos incluirão nossos preditores. Neste artigo, um modelo de previsão que emprega esses preditores será identificado como M_5 .

2.4.3 Outros Preditores

Apesar de apreciarmos a generalidade do Elastic Net, também o comparamos com regressão de Ridge e Lasso, que são aninhados em Elastic Net. A ideia aqui é fazer com que contribuição mais clara em termos de metodologia. Em outras palavras, suponhamos que ambos, Ridge e LASSO, superem as previsões de dicionários fixos; então há valor no uso das metodologias de encolhimento acima e além da metodologia específica escolhida. Acreditamos que isso contribuiria para o artigo, que caso contrário, seria muito dependente da técnica de regularização escolhida. Por esse motivo, consideramos os dois preditores a seguir.

($D_{6,t}$): Este preditor é idêntico a $D_{1,t}$ exceto que usamos LASSO no primeiro estágio para selecionar $X_t^* \in X_t$. Lembre-se de que o LASSO não é robusto para agrupar efeitos e seleciona no máximo T preditores. Neste artigo, um modelo de previsão que emprega esses preditores será rotulado como M_6 .

($D_{7,t}$): Este preditor é idêntico a $D_{1,t}$ exceto que usamos ridge regression no primeiro estágio para selecionar as palavras com maior poder preditivo em X_t . Neste artigo, um modelo de previsão que emprega esses preditores será rotulado como M_7 .

Também consideramos uma análise de componente principal (PCA) pura no conjunto das palavras originais X_t . Este exercício é útil para verificar se o uso de Elastic Net para pré-selecionar as palavras mais preditivas aumenta o poder preditivo do preditor resultante. Assim, consideramos o seguinte preditor.

($D_{8,t}$): Este preditor é idêntico a $D_{1,t}$ exceto que eliminamos o primeiro estágio, ou seja, não usamos Elastic Net(ou LASSO) para pré-selecionar as palavras mais preditivas. Portanto, os preditores baseados em minutas do FED corresponderão a fatores comuns de X_t , em vez de X_t^* . Neste artigo, um modelo de previsão que emprega esses preditores será rotulado como M_8 .

2.5 Análise Empírica

2.5.1 Dados

Obtemos dados trimestrais em tempo real do PIB dos Estados Unidos entre os anos de 1947 e 2017 no Federal Reserve of St. Louis. Então calculamos a taxa de crescimento anualizada do Produto através da fórmula $y_t = \ln(\frac{Y_t}{Y_{t-1}}) * 400$. Esta é a nossa variável dependente e temos 238 observações trimestrais no total. Os dados das atas do FED são agregados em cada trimestre e cobrem o período de 1958(T1)-2017(T2). Embora seja possível obter observações sobre o crescimento do produto e das minutas do FED desde 1947(T2), as atas do FED publicadas antes de 1958(T1) são muito limitadas em termos de comunicação¹⁷. Isso implica que a maioria das palavras encontradas nas minutas do FED depois de 1958T1 quase nunca foram reportadas antes desse período. Assim, para evitar séries temporais sem observações suficientes durante um longo período de tempo, decidimos iniciar nossa amostra em 1958T1.

Dadas 236 observações¹⁸ do crescimento do produto, usamos as primeiras 134 observações, de 1958(T3) ate 1991(T4), como nossa amostra de estimação inicial para os casos com $h = 0$ e $h = 1$. Para horizontes de previsão mais longos $h = 3$ e $h = 6$, as janelas de estimações iniciais são encurtadas, variando de 1958T4 + h a 1992T1 - h . Isso implica que haverá exatamente 102 previsões fora da amostra para cada horizonte de previsão h . Os modelos são estimados recursivamente usando apenas as realizações

¹⁷O número médio de páginas das minutas divulgadas antes de 1958T1 é igual a 20, ao passo que é igual a 60 páginas após 1958(T1).

¹⁸Lembre-se que incluímos 2 defasagens, então a variável dependente inicia em 1958(T3).

em tempo real coletadas das séries, conforme divulgadas em cada trimestre, para gerar previsões em tempo real para cada trimestre. Conforme explicado por [Chauvet and Potter \(2013\)](#), todas as versões históricas não revisadas da série do PIB em tempo real divulgadas a cada mês são coletadas e arquivadas pelo Federal Reserve Bank de Saint Louis e pelo Federal Reserve Bank of Philadelphia. A base de dados trimestral em tempo real usada neste artigo consiste em realizações, ou vintages trimestrais, das séries como teriam aparecido no final de cada trimestre de 1992:T1 a 2017:T2.

Baixamos as atas do FOMC no site do Federal Reserve (FED).¹⁹ Segundo o FED, as atas de reuniões regularmente agendadas são divulgadas antes da divulgação do Produto, como confirmado pela tabela 10, que mostra as datas de divulgação do Produto e da ata do FED correspondente. Esta é uma informação importante porque nos permitirá realizar forecasting e nowcasting. Em outras palavras, podemos usar as informações de minutas do FED disponíveis no momento t para prever o crescimento do produto no tempo $t + h$, $h \geq 0$ onde $h = 0$ corresponde ao que chamamos de nowcast.

O próximo passo é importar as minutas do FED (em formato pdf) para o software R Statistics. O pacote usado para importar essas minutas é o pacote “tm”. O pacote “tm” fornece uma função através da qual é possível importar arquivos pdf para o R²⁰. Com os textos em R, podemos construir um corpus.²¹ Com as minutas do FED no R, podemos construir um corpus e aplicar o pré-processamento dos textos brutos descritos na Seção 2.3.1.

Depois de aplicar os filtros acima, selecionamos, em média, 1.055 palavras por trimestre. Isso cria um problema de alta dimensionalidade na análise de regressão linear e, por essa razão, o método da Elastic Net é empregado. Ao permitir que um número maior de palavras afete a variável dependente, podemos selecionar as mais preditivas sempre que fizermos uma nova previsão. Esta é a ideia principal por trás da abordagem do dicionário que varia no tempo proposta neste artigo.

2.5.2 O procedimento de previsão

[Chauvet and Potter \(2013\)](#) compararam o desempenho de muitos modelos econométricos usados para prever o crescimento do Produto dos EUA de 1992 a 2010. Eles relataram que um modelo simples $AR(2)$ supera modelos mais sofisticados durante períodos de expansão econômica e performa relativamente bem durante as recessões. Suas

¹⁹As atas de 1993 a 2017 podem ser encontradas neste link: <https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm>, e as atas de 1936-1992 podem ser encontradas neste link: https://www.federalreserve.gov/monetarypolicy/fomc_historical_year.htm

²⁰Um rápido tutorial pode ser encontrado nesta página web [Lendo arquivos PDF em R para mineração de texto](#)

²¹A função “Corpus” do pacote “tm” mostra o número de documentos no corpus

descobertas confirmam as evidências iniciais de Nelson (1972) de que previsões a partir de modelos autorregressivos simples são difíceis de superar.

Assim, neste artigo, usamos o seguinte $AR(2)$ linear como um benchmark forte a ser batido, especialmente nos períodos de expansão econômica. Este benchmark corresponde a equação de previsão 2.3 com $W_t = (y_t; y_{t-1})'$ e $\phi_h = 0$, onde y é definido o crescimento do produto em tempo real e h é o horizonte de previsão. Assim, a previsão (MSE) h -passos a frente de y_{t+h} condicional a informação disponível no tempo t , $f_{t+h,t}$, é representada pela média condicional e pode ser computada como: ²²

$$f_{t+h,t} = \hat{c}_h + \hat{\beta}_{1,h}y_t + \hat{\beta}_{2,h}y_{t-1}$$

onde t é a última observação dentro da amostra (ponto de partida da previsão) e $(\hat{c}_h, \hat{\beta}_{1,h}, \hat{\beta}_{2,h})$ são as estimativas MQO de y_{s+h} sobre os preditores $y_s; y_{s-1}$ para $s = 1, \dots, t-h$ com $t = R, \dots, T-h$ (esquema de previsão recursivo). O benchmark $AR(2)$ é aninhado pelos modelos de previsão propostos que são simplesmente a equação (2.3) com $W_t = (y_t; y_{t-1})'$ and $\phi_h \neq 0$. Assim, a previsão (MSE) h -passos a frente de y_{t+h} condicional a informação disponível no tempo t é representada pela média condicional que é:

$$f_{t+h,t}^i = \hat{c}_h + \hat{\beta}_{1,h}y_t + \hat{\beta}_{2,h}y_{t-1} + \hat{\phi}_h' D_{i,t}, \quad (2.7)$$

onde $D_{i,t}$ $i = 1, \dots, 8$ correspondem a um dos 8 conjunto de preditores gerados na seção anterior e $(\hat{c}_h, \hat{\beta}_{1,h}, \hat{\beta}_{2,h}, \hat{\phi}_{2,h})$ são as estimativas MQO de y_{s+h} sobre os preditores $(y_s, y_{s-1}, D_{i,s})$ para $s = 1, \dots, t-h$ com $t = R, \dots, T-h$. Assim, se as minutas do FED tem algum poder preditivo, então a previsão correspondente fora da amostra deve superar o benchmark $AR(2)$.

Se os dicionários variantes no tempo forem mais preditivos do que os fixos, as previsões fora da amostra do modelo do modelo (2.7) com $D_{i,t}$ $i = 1, 2, 3, 6, 7$ devem superar as previsões do modelo (2.7) com $D_{i,t}$ $i = 4, 5$. Se o elastic net funciona melhor na seleção das palavras, então o modelo com 2.7 $D_{i,t}$ $i = 1$ deve superar as previsões fora da amostra a partir do modelo (2.7) com $D_{i,t}$ $i = 6, 7$. Finalmente, se a pré seleção das palavras ajuda a aumentar a acurácia da previsão, então o modelo (2.7) com $D_{i,t}$ $i = 1, 6, 7$ deve superar as previsões a partir do modelo (2.7) com $D_{i,t}$ $i = 8$.²³ Assim, esse

²²Ao longo de todo este artigo assumimos que a variável dependente y_t é um processo covariância estacionário.

²³Relembre que os modelos $D_{i,t}$ $i = 1, 6, 7$ usam o mesmo procedimento para computar os preditores finais, eles apenas diferem sobre o método de aprendizado de máquina usado para selecionar as palavras mais preditivas.

arcabouço é rico o suficiente para investigar não apenas se as minutas do FED contêm informações preditivas fortes, mas também se tal poder preditivo depende de ambos, o uso do dicionário variante no tempo e o método usado para selecionar as palavras mais preditivas.

Uma vez que o crescimento real do PIB no tempo t é divulgado após as atas do FED, podemos usar a minuta do FED disponível no momento t para prever o crescimento do PIB em t , ou seja, nós podemos calcular

$$f_{t,t}^i = \hat{c} + \hat{\beta}_1 y_{t-1} + \hat{\beta}_2 y_{t-2} + \hat{\phi}' D_{i,t}, \quad (2.8)$$

onde $(\hat{c}, \hat{\beta}_1, \hat{\beta}_2, \hat{\phi}')$ são estimativas MQO a partir da regressão da série de tempo de y_s sobre y_{s-1} , y_{s-2} e $D_{i,s}$; $f_{t,t}^i$ é o nowcast em tempo real do crescimento do Produto que depende de y_{t-1} e y_{t-2} bem como do valor do(s) preditor(es) $D_{i,t}$ na origem da previsão t . Assim, o nowcast é uma previsão h-passos a frente com $h = 0$.

Como citado por [Chauvet and Potter \(2013\)](#), uma previsão de julgamento ocorre quando as previsões obtidas a partir de modelos econométricos são ajustadas por seus usuários. Alguns dos julgamentos se baseiam em informações subjetivas extraídas das atas do FED e/ou outras fontes de informação sobre o curso futuro da economia. Como a nossa abordagem depende de informações das minutas do FED para melhorar as previsões de um modelo $AR(2)$, poderíamos esperar que as previsões a partir de pesquisas tal como o Blue Chip²⁴, poderiam ter informação adicional que melhoraria a previsão fora da amostra do crescimento do Produto. Testamos tal hipótese na seção empírica deste artigo. Assim, além dos modelos de previsões propostos neste artigo, também incluímos a previsão de julgamento a partir dos indicadores Blue Chip (BC). Consideramos as previsões do Blue Chip individualmente mas também adicionamos ela a um modelo $AR(2)$, dando origem ao que chamamos modelo $AR(2)$ - BC .²⁵ Finalmente, consideramos um modelo de combinação de previsão(FC) com pesos iguais de todos os modelos baseados nas minutas do FED, que compreende as previsões dos modelos $M_i, i = 1, \dots, 8$. Na totalidade, nosso conjunto de modelos de previsão incluem $f_{t+h,t}^j \in (M_i, BC, AR(2)-BC, FC)$

²⁴Os indicadores Blue Chip representam uma pesquisa no qual participam em torno dos 50 melhores economistas de Bancos, Indústria, Corretoras e Seguradoras. A pesquisa vem sendo realizada desde 1976 e compreende várias séries macroeconômicas, incluindo o crescimento do PIB.

²⁵ Existe uma previsão do Blue Chip para cada horizonte de previsão considerado neste artigo. Além do mais as previsões BC são divulgadas mais de uma vez no trimestre. Sempre consideramos a última divulgação do trimestre, o que tipicamente ocorre 10 dias após o fim do trimestre.

2.5.3 Avaliação da Previsão

Para avaliar a performance dos diferentes modelos de previsão da taxa de crescimento do *PIB*, computamos a raiz do erro quadrático médio (*RMSFE*) de cada modelo $f_{t+h,t}^i$, $i = 1, \dots, 8$, relativo ao benchmark *AR*(2), $f_{t+h,t}$. O *RMSFE* é calculado como:

$$RMSFE_j^h = \frac{\sqrt{\sum_{t=1}^P (y_{t+h} - f_{t+h,t}^j)^2}}{\sqrt{\sum_{t=1}^P (y_{t+h} - f_{t+h,t})^2}} \quad (2.9)$$

onde P é o número das previsões h -passos a frente fora da amostra (*OOS*), $h \geq 0$. Se o valor de $RMSFE_i$ é menor que 1, então o modelo i performa melhor que o modelo *AR*(2) em termos de *RMSFE*, produzindo melhor previsão(nowcasts) do Produto.

Para testar se um modelo condicional i produz previsões significativamente melhores, testamos a hipótese nula de que modelos produzem previsões iguais estatisticamente proposta por [Clark and West \(2007\)](#). Escolhemos este teste porque o benchmark *AR*(2) é aninhado por todos os outros modelos de previsão.²⁶ Para erros de previsão autorrelacionados, ($h > 1$) usamos [Newey and West \(1986\)](#) para estimar os desvios padrões consistentes com a autocorrelação. Finalmente, uma descoberta recente de [Gonçalves et al. \(2017\)](#) justifica usar os valores críticos usuais no teste de igualdade das previsões com modelos que usam fatores.

É importante mencionar novamente que não pretendemos propor um modelo de previsão última para o crescimento do Produto mas sim combinar novos desenvolvimentos a partir da regressão de texto e Aprendizado de Máquina Supervisionado em uma maneira que é inédita para identificar a informação mais preditiva das minutas do FED. Somos conscientes que outros modelos de previsão, tal como os que incluem uma estrutura de lags D_t nas equações de previsão (2.7 and 2.8), poderiam gerar previsões mais acuradas para o crescimento do produto em tempo real. Além do mais, tais preditores também poderiam ser usados para aumentar a precisão dos modelos de previsão existentes, como os modelos MIDAS propostos por [Andreou et al. \(2013\)](#), [Carriero et al. \(2015\)](#) e [Marcellino et al. \(2016\)](#). Nesse caso, estaríamos explorando o poder de previsibilidade das informações de alta frequência e de texto para a previsão do crescimento do produto. Apresentamos nossos principais resultados na próxima seção.

²⁶O teste por [Diebold and Mariano \(2002\)](#) foi projetado para comparar modelos não aninhados. Se os modelos de previsão estiverem aninhados, então o teste DM pode ser subdimensionado sob a hipótese nula e pode ter baixo poder sob a hipótese alternativa.

2.6 Resultados

Nesta seção, reportamos nossos resultados para o período completo fora da amostra, 1992(T1)-2017(T2), bem como para dois sub-períodos. O primeiro inclui apenas trimestres para os quais o crescimento real do PIB foi negativo: usamos a classificação de recessão do NBER, da qual os períodos 2001T1-2001T4 e 2007T4-2009T2 foram classificados como recessões; o segundo sub-período inclui trimestres para os quais o crescimento do PIB foi positivo (período de expansão). De acordo com o [Chauvet and Potter \(2013\)](#), as previsões de $AR(2)$ são especialmente boas para prever o crescimento do produto durante os períodos de expansão. Assim, para concluir que as informações das atas do FED contêm poder preditivo, os modelos de previsão propostos devem superar o benchmark $AR(2)$ não apenas quando ele está não performa bem (períodos de recessão), mas também quando performa melhor (períodos de expansão).

A tabela 11 exibe os valores do RMSFE (equação 2.9) em relação ao benchmark $AR(2)$, bem como o resultado do teste CW para a hipótese nula de modelos de igual precisão. As previsões baseadas em dicionários fixos (modelos 4 e 5) não parecem adicionar muito poder preditivo ao benchmark $AR(2)$, já que seu RMSFE relativo não é estatisticamente diferente de 1. Baseado neste resultado, conclui-se que as informações contidas nas atas do FED não ajudam a prever o crescimento do produto. No entanto, a principal contribuição deste artigo é mostrar que os dicionários fixos não nos permitem extrair o melhor das minutas do FED, uma vez que o vocabulário usado pelos bancos centrais provavelmente está mudando ao longo do tempo. De fato, a tabela 11 também mostra que as previsões baseadas no dicionário que variam no tempo ($f_{t+h,t}^i$, $i = 1, 2, 3, 7$.) superam o benchmark $AR(2)$ por uma boa margem e o resultado do teste CW indica a rejeição da hipótese nula de precisão igual ao nível usual de 10% de significância. O modelo M_1 se destaca como a melhor previsão baseado nas minutas do FED.

Embora o uso de LASSO e ridge regression para selecionar as palavras mais preditivas (modelos M_6 e M_7) melhore a precisão da previsão em relação ao benchmark, elas não melhoram em relação ao Elastic Net(modelo M_1). Assim, permitir um dicionário que varie no tempo melhora a precisão da previsão, mas o método empregado para selecionar as palavras mais preditivas também é importante. Os resultados para o modelo M_8 mostram a importância de pré-selecionar as palavras mais preditivas antes de computar os fatores comuns. De fato, tabela 11 mostra que os modelos M_i , $i = 1, 2, 3, 7$ superam o modelo M_8 , sugerindo que a pré-seleção de palavras ajuda a aumentar a precisão da previsão. Finalmente, as previsões Blue Chip (BC) e $AR(2)$ -BC não têm um bom desempenho relativamente ao benchmark, sugerindo que, se formos capazes de construir dicionários variando no tempo, as informações extraídas das minutas do FED podem têm mais poder de previsão do que as informações usadas pelos analistas individuais.

Esses achados representam a primeira evidência empírica sobre o conteúdo preditivo das atas do FED e ressaltam a importância de combinar técnicas de aprendizado de máquina (Elastic Net) com dados textuais para melhorar o desempenho dos modelos de previsão macroeconômicos.

A tabela 12 mostra que alguns resultados notáveis são observados quando dividimos o período completo fora da amostra nas fases dos ciclos de negócios. O desempenho de todos os modelos é melhor na previsão de expansões do que nas recessões. Todos os modelos são menos precisos para prever o crescimento do produto durante as recessões, mas essa perda é menor nos modelos que usam os preditores baseados em minutas do FED calculados a partir de dicionários que variam no tempo ($f_{t+h,t}^i$, $i = 1, 2, 3, 6, 7$). Esse resultado sugere que as informações das minutas do FED também podem ser usadas para minimizar a perda de precisão dos modelos de previsão durante os períodos de recessão, desde que permita dicionários que variem no tempo. Observe que os modelos que usam (LASSO e ridge regression) (M_6 e M_7) não superam M_1 durante qualquer período, seja recessão ou expansão, sugerindo novamente que o método usado para pré-selecionar as palavras também importa bastante. Tabelas 11 e 12 também sugerem que combinando as previsões dos modelos que incorporam informações das minutas do FED, M_i , $i = 1, 2, \dots, 8$, produzem ganhos adequados de precisão relativos ao modelo benchmark $AR(2)$, especialmente durante os períodos de expansão. A análise do teste CW também sugere que uma simples combinação de previsão com pesos iguais não pode ser considerada estatisticamente pior do que o benchmark $AR(2)$. Estes resultados suportam a conclusão de que se formos capazes de extrair a informação mais preditiva das minutas, então as previsões baseadas nas minutas do FED superarão o benchmark $AR(2)$ exatamente durante os períodos em que ele tem sido imbatível.

Finalmente, a tabela 13 relata os resultados para o nosso exercício de nowcasting. A principal motivação para este exercício baseia-se no fato de que o crescimento real do PIB no trimestre t é divulgado após as atas. Lembre-se de que um nowcast é calculado usando a equação (2.8) e o benchmark do nowcast usa a mesma equação sem os preditores $D_{i,t}$. A tabela 13 confirma que as minutas do FED contêm poder preditivo quando se considera um dicionário que varia no tempo. Também descobrimos que a combinação de peso iguais dos modelos de previsão funciona muito bem.

Os resultados relatados nesta seção dão suporte à conclusão de que, se permitirmos o vocabulário variante no tempo, as informações do texto nas atas do FED podem ser usadas para prever o crescimento da produção. Na próxima seção, explicamos por que a abordagem de previsão proposta funciona tão bem.

{ Tabela 11 aqui. }

{ Tabela 12 aqui. }

{ Tabela 13 aqui. }

{ Tabela 14 aqui. }

2.6.1 Explicando o sucesso das previsões baseadas nas minutas do FED

Nesta seção, decompomos o erro quadrático médio ($MSFE$) em duas partes: variância e viés. Calculamos $MSFE$ para qualquer previsão $f_{t+h,t}^j \in (M_i, BC, AR(2)-BC, FC)$ como $\frac{1}{P} \sum_t (y_{t+h} - f_{t+h,t}^j)^2$ e a variância incondicional $\frac{1}{P} \sum_t \left(f_{t+h,t}^j - \frac{1}{P} \sum_t f_{t+h,t}^j \right)^2$, onde P é o número total de observações fora da amostra. O viés ao quadrado é calculado como a diferença entre o $MSFE$ e a variância da previsão (Elliott and Timmermann, 2013b, Rapach and Strauss, 2010) e Lima and Meng (2017).

Figuras 14 - 17 descrevem a variância de previsão relativa e viés de previsão ao quadrado de todos os modelos de previsão $f_{t+h,t}^i$, rotulado como M_i $i = 1, \dots, 8$; Blue Chip(BC); AR2-BC e combinação de previsão com pesos iguais(FC) para o período completo fora da amostra. A variância de previsão relativa (viés ao quadrado) é calculada como diferença entre a variância da previsão (viés ao quadrado) do j -ésimo modelo e a variância de previsão (viés ao quadrado) do modelo $AR(2)$. Consequentemente, o valor da variância de previsão relativa (viés ao quadrado) para o $AR(2)$ é necessariamente igual a zero. Cada ponto da linha sólida representa uma previsão com o mesmo $MSFE$ que o $AR(2)$; pontos à direita da linha estão as previsões superadas pelo $AR(2)$, e pontos à esquerda representam previsões que superam o $AR(2)$. Finalmente, tanto a variância como o viés ao quadrado são medidos na mesma escala, de modo que é possível determinar o trade-off entre variância e viés de cada modelo de previsão.

Como o modelo $AR(2)$ é uma versão parcimoniosa (aninhada) dos modelos de previsão baseados em texto, as previsões $f_{t+h,t}^i$, $AR(2)$ terão a menor variância entre todas as outras previsões, mas será viesado. Portanto, um modelo de previsão $f_{t+h,t}^i$ superará o benchmark $AR(2)$ apenas se for capaz de reduzir substancialmente o viés de previsão relativamente ao $AR(2)$. Figuras 14 - 17 mostra a decomposição ($MSFE$) para vários horizontes de previsão $h = 1, 3, 6$, e nowcasting ($h = 0$). De acordo com nossa explicação, se alguém usa Elastic Net para selecionar as palavras mais preditivas ao longo do tempo, então podemos esperar que $f_{t+h,t}^i$, $i = 1, 2, 3$ minimizará o viés e, consequentemente, superará o benchmark.

Figuras 14 - 17 confirmam que a previsão do benchmark $AR(2)$ tem a menor variância, mas é tendenciosa. Portanto, o único canal pelo qual um modelo de previsão

pode superar o benchmark é através de uma redução do viés de previsão. Um bom modelo será aquele que pode reduzir substancialmente o viés enquanto mantém a variação sob controle. Os resultados exibidos nessas figuras mostram que os modelos de previsão que dependem de dicionários fixos, $f_{t+h,t}^i$, $i = 4, 5$ são incapazes de reduzir substancialmente o viés, mantendo a variância das previsões sob controle e, por essa razão, eles não superam o benchmark $AR(2)$ por uma grande margem. Por outro lado, podemos ver que o bom desempenho dos modelos de previsão que dependem de dicionários variantes no tempo, $f_{t+h,t}^i$, $i = 1, 2, 3, 6, 7$, decorre de sua capacidade de reduzir substancialmente o viés ao quadrado da previsão sem aumentar muito a variância de previsão. Figuras 14 - 17 mostram que esta descoberta ocorre em todos os horizontes de previsão.

A discussão anterior oferece uma explicação para os resultados apresentados na seção anterior. As informações contidas nas atas do FED são úteis para prever o crescimento do PIB porque reduzem substancialmente o viés da previsão, em relação ao benchmark $AR(2)$, sem aumentar muito a variância da previsão. Esse resultado depende da nossa capacidade de usar aprendizado de máquina (Elastic Net) para identificar as palavras mais preditivas das atas do FED. Também nos ajuda a entender por que a abordagem popular de regressão de texto baseada em dicionários fixos não é útil para a previsão macroeconômica. De fato, os modelos de previsão baseados em dicionários fixos não reduzem substancialmente o viés da previsão e, por essa razão, dificilmente ou nunca superam o benchmark $AR(2)$.

A tabela (15) nos ajuda a entender por que um dicionário variante no tempo deve ser interpretado como um dispositivo poderoso para previsão macroeconômica. Ele exibe uma lista com as cinco palavras mais preditivas durante o período fora da amostra. Especificamente, Tabela (15) mostra a frequência média de classificação de grupos de palavras identificados por sua raiz ou termo. Essa classificação é baseada no valor absoluto dos coeficientes em X_t , estimado por Elastic Net. Os dez maiores coeficientes são usados para identificar as 5 palavras principais.²⁷ A primeira conclusão desta tabela é que algumas palavras como “economic activity” têm poder de previsão durante os períodos de recessão, mas é selecionada com menor frequência durante os tempos de expansão. Outras palavras como “oil” foram menos selecionadas durante os períodos de recessão. Essa análise sugere que as informações contidas nas atas do FED são de fato úteis para a previsão macroeconômica, mas o poder preditivo de algumas palavras é bastante diferente em períodos distintos. Assim, o sucesso de um modelo de previsão que utiliza informações de minutas do FED depende de sua capacidade de identificar as palavras mais preditivas ao longo do tempo, e é exatamente isso que é feito pelo método proposto neste artigo.

²⁷Este método de classificação funciona porque todas as séries temporais X_t são medidas na mesma escala normal padrão.

{ Tabela 15 aqui. }

2.6.2 Previsões do Livro verde do FED e previsões baseadas nas minutas do FED

Uma questão natural é se as atas do FED têm valor acima e além da previsão da equipe econômica do FED(previsão do Livro Verde). Há várias razões para concluir que as atas do FED fornecem informações adicionais úteis para a previsão macroeconômica. A fim de elaborar sobre este ponto, descrevemos abaixo algumas etapas de uma reunião do FOMC(Federal Open Market Committee):²⁸

estágio i. Um funcionário do FED de Nova York apresenta desenvolvimentos do mercado financeiro e de câmbio, e a equipe responde a perguntas sobre essas condições financeiras

estágio ii. Há uma discussão sobre situação econômica:

- a. A equipe do Conselho de Governadores apresenta as previsões do Livro Verde;
- b. Há uma série de perguntas sobre a apresentação do staff do FED;
- c. Os membros do FOMC apresentam sua opinião sobre as perspectivas econômicas.

Essas discussões trazem um novo conjunto de informações(opiniões e comentários) que ainda não foi incorporado nas previsões do Livro Verde. Há também uma discussão de estratégia de política monetária, onde mais informações são apresentadas.

etapa iii. Discussão da estratégia de política monetária:

- a. O diretor de assuntos monetários do conselho apresenta uma variedade de alternativas de política monetária;
- b. uma rodada potencial de perguntas do Staff;
- c. o presidente e os outros membros do FOMC discutem suas preferências de políticas.
- v. O FOMC vota na decisão de política.

Como podemos ver na descrição acima, novas informações são geradas por meio de discussões que ocorrem durante uma reunião do FOMC. Para ajudar os escritores das minutas, as reuniões do FOMC são gravadas e as fitas(gravações) são usadas mais tarde

²⁸extraímos essas informações de Hansen et al. (2017)

como uma fonte de informação para escrever uma minuta. Se tal informação for útil para a previsão econômica, então as previsões baseadas nas minutas do FED propostas neste artigo devem superar as previsões do Livro Verde. Tabela 16 relata o RMSFE relativo ao modelo de benchmark AR(2). Reportamos o desempenho fora da amostra do Livro Verde e as previsões obtidas usando o modelo M_1 para $h = 0, 1, 3, 6$.²⁹ Observe que a previsão do Livro Verde é publicada apenas com um atraso de 5 anos e, por esse motivo, terminamos nossa análise fora de amostra em 2012Q4.³⁰ A principal conclusão da Tabela 16 é que tanto o livro verde quanto o modelo M_1 superam o benchmark AR(2) para $h = 0$ (nowcasting), mas somente a previsão baseada em minutas do FED (M_1) supera o benchmark por uma grande margem para horizontes mais longos. Consideramos esse resultado como uma forte evidência de que as informações geradas durante uma reunião do FOMC são úteis para previsões fora da amostra.

2.6.3 Análise da Função Impulso Resposta

Nesta seção, mostramos algumas evidências empíricas do mecanismo de transmissão dos choques de comunicação do FED para valores futuros do yield spread. Este achado empírico é identificado através da estimação de uma função estrutural de impulso resposta (Lucca and Trebbi, 2009). Calculamos o yield spread como a diferença entre o rendimento dos títulos americanos de 10 anos e 3 meses coletados do banco de dados do FRED disponível no FED-Saint Louis (T10Y3M). Escolhemos esses dois vencimentos de títulos porque além de serem os vencimentos de curto e longo prazo mais acompanhados pelo mercado financeiro, o estudo realizado por Hansen and Seo (2002) demonstra que as séries tem uma relação de longo prazo significativa.

Estimamos um modelo VAR com variáveis ordenadas como {fator hawkish, fator dovish, spread}, onde o fator hawkish (dovish) é o fator comum para os bigramas hawkish (dovish) selecionados pela abordagem M_3 . Lembre-se de que esses fatores só carregam informações hawkish (dovish) e, portanto, podem ser usados em nossa análise de impulso resposta para identificar o efeito de um choque na comunicação hawkish (dovish) do FED sobre os valores futuros do yield spread. Esperamos que os choques no fator hawkish aumentem o yield spread (relação positiva), enquanto os choques no fator dovish diminuam o yield spread (relação negativa).

A fim de estimar da função impulso-resposta estrutural, aplicamos uma restrição de curto prazo (fatoração de Choleski) onde, dado o ordenamento das variáveis no modelo VAR, o efeito contemporâneo do yield spread sobre os dois fatores é zero. Os dados sobre

²⁹Consideramos apenas as previsões do Livro Verde da última reunião de cada trimestre.

³⁰Isso explica por que os resultados relatados para o M_1 na Tabela 16 são diferentes dos reportados anteriormente em Tabelas 11 e 13.

o yield spread iniciam em 1984(T1) e terminam em 2017(T2). Também multiplicamos o yield spread por 100, de modo que a resposta a um choque de um desvio padrão nos fatores seja expressa em pontos base. Finalmente, o modelo VAR foi estimado com 2 lags (critério de informação de Schwarz) e usamos bootstrap com 100 simulações para estimar um intervalo de confiança para a função impulso-resposta.

As figuras (18) e (19) mostram as respostas do yield spread a choques inesperados de um desvio padrão nos fatores hawkish e dovish. Um choque positivo no fator hawkish é subsequentemente seguido por um aumento no spread. O spread exibe uma resposta em forma de corcunda a um choque no fator hawkish com um pico de cerca de 15 pontos base após 5 trimestres. Por outro lado, um choque positivo no fator dovish (Figura (19) é seguido por uma diminuição do spread com um pico de cerca de 25 pontos base após 5 trimestres. As linhas vermelhas mostradas na Figura (18) e (19) são os intervalos de confiança de dois desvios padrões. A principal conclusão desta análise de impulso resposta é que os choques na comunicação do FED são transmitidos para o yield spread com um pico desta transmissão ocorrendo em 5 trimestres. Essa descoberta nos ajuda a entender por que os preditores baseados em minutas do FED propostos nesta tese são úteis para previsões macroeconômicas.

{ Figura 18 aqui. }

{ Figura 19 aqui. }

2.7 Considerações Finais

Palavras importam. As comunicações do banco central podem indicar as direções esperadas da política monetária da economia. A autoridade monetária com forte credibilidade é capaz de controlar as expectativas dos agentes privados. Palavras usadas em minutas do FED, portanto, têm poder de mercado. O canal através do qual este resultado opera é que choques positivos à comunicação hawkish(dovish) levam a um aumento (diminuição) no yield spread e, portanto, são interpretados como choques de política monetária contracionistas(expansionistas). Uma simples análise de impulso resposta conduzida neste artigo empresta suporte empírico a este canal.

Pesquisadores aplicados, formuladores de políticas e vários investidores estão sempre em busca de indicadores úteis para prever com sucesso futuras atividades econômicas. A literatura de previsão mostrou que um modelo simples e parcimonioso de séries temporais $AR(2)$ frequentemente supera modelos macroeconômicos mais complexos na previsão de crescimento econômico, especialmente durante períodos de expansões econômicas.

Curiosamente, propomos o aumento do modelo simples $AR(2)$ incluindo preditores computados de um dicionário que varia com o tempo e descobrimos que esse novo modelo possui poder preditivo superior. Este não foi o caso quando usamos preditores construídos a partir de dicionários fixos. Como explicado anteriormente, o dicionário variante no tempo é mais capaz de captar o tom do banco central a qualquer momento e, portanto, pode ser usado para construir melhores preditores das atividades econômicas futuras.

Nossos resultados sugerem que a combinação de aprendizado de máquina (Elastic Net) e regressão textual tem potencial para identificar preditores baseados em documentos que são mais poderosos do que a contrapartida construída a partir de dicionários fixos. Nossas descobertas mostram que as previsões baseadas em tais preditores são capazes de reduzir substancialmente o viés de previsão em relação ao benchmark $AR(2)$ enquanto mantêm a variância da previsão sob controle. Assim, tem implicações significativas para pesquisadores empíricos na avaliação dos efeitos da política monetária sobre a economia. A identificação de choques na política monetária continua a ser um desafio importante entre os pesquisadores. Os preditores identificados em um cenário de tempo variável, como mostramos neste estudo, poderiam potencialmente capturar verdadeiros choques de política monetária. Pesquisas futuras devem explorar este canal para avaliar os efeitos da política monetária sobre a produção e outros macro-agregados importantes.

Capítulo 3

Previsão do Prêmio de Risco: Identificando o Poder Preditivo das Notícias Financeiras

3.1 Introdução

A previsão de retornos de ações e prêmio de risco tem sido um quebra-cabeça na literatura de econometria financeira. Variáveis financeiras são consideradas de difícil previsibilidade, principalmente em frequências mais altas como mensal, diária e intra diária. Analistas financeiros, fundos de investimento, investidores individuais dependem de previsões do índice do mercado de ações na sua tomada de decisão. Muitos estudos mostram que variáveis econômicas e/ou contábeis, como por exemplo a taxa de dividendo, spread, yield dos títulos do tesouro, entre outras pode ter algum poder preditivo durante dado período de tempo, mas em outras épocas são considerados preditores fracos. No entanto, uma das questões que impulsionam a busca de novos preditores e a possibilidade dos mesmos preverem retornos negativos, dado que a média histórica, na maioria dos casos quando se usa uma amostra longa, trará a previsão apenas de retornos positivos. Com um modelo que preveja com acurácia além de retornos positivos, retornos negativos, o tomador de decisão pode naquele determinado período zerar sua posição em renda variável ou até mesmo abrir uma posição de venda, seja ela via mercado futuro ou ETFs inversos¹. Assim, uma das motivações do artigo é a busca de preditores de dados estruturados ou não estruturados que melhorem a previsão do prêmio de risco.

¹Por exemplo, ao comprar ETF negociado com o ticker “SPXS” negociado na NYSE, o investidor ganhará com a queda do índice S&P 500.

Temos vários artigos que estudam esse assunto e tentam descobrir novos métodos e/ou preditores que aumentam a precisão da previsão do prêmio de risco, como [Welch and Goyal \(2007\)](#), [Rapach and Strauss \(2010\)](#). [Lima and Meng \(2017\)](#) entre outros. [Welch and Goyal \(2007\)](#) reexaminam o desempenho de variáveis que foram sugeridas pela literatura acadêmica como bons preditores do prêmio de risco e acham que, em geral, esses modelos têm um mal desempenho tanto dentro da quanto fora da amostra. [Rapach and Strauss \(2010\)](#) empregam algumas estratégias para melhorar a precisão das previsões. Eles concluem que essas estratégias, como incluir restrições no modelo motivadas economicamente, combinação de previsão, índices de difusão e mudanças de regime melhoram o desempenho da previsão. Os autores argumentam que a incerteza e a instabilidade substancial do modelo na estimação dos parâmetros prejudica seriamente a capacidade de previsão dos modelos de previsão que usam preditores individuais.

[Lima and Meng \(2017\)](#) desenvolvem um novo método de previsão que minimiza os efeitos de preditores fracos e erros de estimação sobre a precisão das previsões de prêmio de risco. O método proposto baseia-se num esquema do cálculo da média aplicado a quantis condicionais em preditores selecionados pelo LASSO. Eles acham que as previsões resultantes superam a média histórica, e outros modelos existentes, por margens estatisticamente e economicamente significativas. [Rapach and Zhou \(2013\)](#) utilizam dados mensais em uma amostra entre 1957 e 2010 e encontram que o modelo sum-of-the-parts(SOP) supera a média histórica por uma margem significativa. Esse modelo proposto originalmente por [Ferreira and Santa-Clara \(2011\)](#), baseia-se em uma estimação por Mínimos Quadrados restrito e tem como preditores o índice dividendo sobre o preço (DP) e a média móvel de 20 anos da taxa de crescimento do lucro por ação. Nesse caso o intercepto é igual a média da taxa de crescimento do lucro menos a taxa livre de risco e o parâmetro da variável DP é restrito para ser igual a 1.

Em um artigo seminal, [Garcia \(2013\)](#) estuda o efeito do sentimento sobre os preços dos ativos durante o século XX (1905 a 2005). O autor usa textos do The New York Times para construir variáveis de sentimento usando análise textual. Ele conclui que a previsibilidade dos retornos das ações usando o conteúdo de notícias é particularmente forte durante as recessões. Ao contrário de [Garcia \(2013\)](#), que usa um dicionário fixo criado por [Loughran and McDonald \(2015\)](#), usamos uma abordagem de dicionário variante no tempo baseada em [Lima et al. \(2018b\)](#). Outra diferença entre o nosso trabalho e o artigo de [Garcia \(2013\)](#) é que realizamos uma análise fora de amostra.

[Rapach et al. \(2016\)](#) mostram que a variável short interest² é indiscutivelmente o mais forte preditor para o retorno de ações. De acordo com [Rapach et al. \(2016\)](#)

²Variável que indica o número de ações vendidas à descoberto em relação ao total de ações em circulação

ele supera uma série de preditores de retornos populares dentro e fora da amostra, com estatísticas anuais de R^2 de 12,89% e 13,24%, respectivamente, para $h=12$. Os resultados encontrados pelos autores indicam que a variável short interest pode gerar ganhos de utilidade de mais de 300 pontos-base por ano para um investidor de média variância. A análise feita via vetor auto regressivo pelo artigo mostra que a fonte do poder preditivo da variável short interest vem do canal do fluxo de caixa e que os short sellers são traders informados que são capazes de antecipar o fluxo de caixa e consequentemente os retornos associados.

[Peress \(2014\)](#) investigou o impacto causal na negociação e na formação de preços dos ativos, examinando as greves de jornais em vários países. O volume de negócios caiu 12% nos dias de greve e a dispersão dos retornos das ações e sua volatilidade intra-diária foram reduzidas em 7%, enquanto retornos agregados não foram afetados. Além disso, a análise do poder de previsibilidade do retorno indica que os jornais propagam notícias do dia anterior. O autor demonstra que a mídia contribui para a eficiência do mercado de ações, melhorando a disseminação de informações entre os investidores e sua incorporação nos preços das ações.

[Høghaug Larsen and Thorsrud \(2017\)](#) decompõem os dados textuais diários de um jornal de negócios norueguês em tópicos de notícias e investigam seu papel preditivo e causal para os preços dos ativos. As três principais descobertas dos autores são: (1) uma inovação de um desvio padrão nos tópicos de notícias prevê aproximadamente um aumento de 1 ponto percentual nos retornos da abertura do mercado em relação ao fechamento e padrões de continuação significativos com um pico de 4 pontos percentuais após 15 dias úteis, com poucos sinais de reversão; (2) estratégias simples de investimento baseadas em notícias geram retornos ajustados ao risco anualizados significativos de até 20%; e (3) durante uma escassez de mídia, devido a uma greve exógena, os retornos para empresas particularmente expostas à medida de notícia sofrem uma queda substancial. As estimativas dos autores sugerem que entre 20% e 40% do poder preditivo dos tópicos de notícias se deve ao efeito causal da mídia. Juntas, essas descobertas dão forte apoio a uma visão de atenção racional, em que a mídia alivia fricções de informação e dissemina informações fundamentais para uma grande população de investidores.

[Heston and Sinha \(2015\)](#) usam um conjunto de dados de mais de 900.000 notícias para testar se as notícias podem prever retornos de ações. Os autores medem o sentimento com uma rede neural proprietária da Thomson-Reuters e descobriram que as notícias diárias preveem retornos de ações por apenas 1 a 2 dias, confirmando pesquisas anteriores. Notícias semanais, no entanto, preveem retornos de ações para um trimestre. Notícias positivas aumentam rapidamente os retornos das ações, mas as reportagens negativas

têm uma reação demorada. Grande parte da resposta tardia às notícias ocorre em torno do anúncio dos resultados subsequentes.

Neste artigo, usamos os preditores baseados em notícias financeiras para prever o prêmio de risco. Os agentes do mercado financeiro contam com previsões de prêmio de risco para tomar decisões sobre alocação de portfólio. Uma das previsões mais utilizadas, a média histórica, é uma previsão que, apesar de superar muitos modelos que usam preditores tradicionais, não traz nenhuma informação para o investidor, pois para o mercado americano por exemplo, a média histórica é positiva a longo prazo. A utilidade em se usar as informações das notícias financeiras, bem como outros preditores é que os mesmos são capazes de prever retorno negativo e, para esse determinado período, mostrar ao investidor que a melhor tomada de decisão seria investir apenas em ativos livre de risco naquele em período, zerando a posição em ativos de risco.

Nossos resultados apontam que a utilização das notícias financeiras traz ganhos tanto estatístico quanto econômico para os investidores. Um das fontes desse ganho, é que os modelos baseados em notícias tem uma boa performance em recessões, em que temos uma maior probabilidade de ocorrência de retornos negativos. Também outra fonte de ganho ocorre nas expansões, pois os modelos de notícias em sua maioria tem um bom desempenho nesses sub períodos. Uma das fontes dos ganhos econômicos na expansão é que os modelos basados em notícias, ao detectarem que o sentimento está muito positivo, conseguem prever um prêmio de risco bem maior que a média histórica. Para a nossa utilidade média-variância adotada, em que é permitido alavancagem no ativo de risco, o modelo baseado em notícias passa a mensagem ao investidor que para aquele determinado período ele pode ser expor ao máximo em ativos de risco.

O restante deste artigo está organizado da seguinte forma. A seção 3.2 apresenta os dados utilizados pela pesquisa, os procedimento de construção dos preditores baseados em notícias financeiras e a avaliação da previsão. A seção 3.3 discute os resultados e apresenta as ganhos tanto em termos estatísticos como econômicos trazidos pela utilização das informações contidas nas notícias. Por fim, a seção 3.4 apresenta as considerações finais e sugestões para futuras pesquisas na área.

3.2 Dados, procedimentos e avaliação da previsão

3.2.1 Dados

Nossos dados ³ contêm observações mensais do prêmio de risco para o índice S&P 500 r_t e 15 preditores, que incluem razão de preço-dividendo (DP), dividend yield (DY), razão lucro-preço (EP), taxa de pagamento de dividendos (DE), volatilidade do índice (SVAR), índice book-to-market (BM), expansão do patrimônio líquido (NTIS), taxa de juros do Tesouro (TBL), rendimento a longo prazo (LTY), retorno a longo prazo (LTR), term spread(TMS), default yield spread(DFY), yield spread(DFR), inflação (INFL) e média móvel da razão lucro/preço (E10P), de janeiro de 1980 a dezembro de 2017. Ao contrário de [Welch and Goyal \(2007\)](#) e seguindo [Lima and Meng \(2017\)](#), não defasamos o preditor INFL, o que implica que estamos assumindo expectativas adaptativas para futuras mudanças de preços. Os preditores econômicos serão denotados por X_t . Nossa amostra começa em janeiro de 1980 devido à disponibilidade de notícias financeiras no Banco de Dados Dow Jones Factiva.

Nós baixamos as notícias financeiras publicadas no “The Wall Street Journal” e no “The New York Times” da base de dados Dow Jones Factiva de 1980:1 a 2017:12. Nesse caso, definimos nosso Corpus no tempo t como todas as notícias coletadas naquele determinado mês. Para selecionarmos apenas as notícias financeiras, filtramos pelos temas “Economic News” e “Commodity/Financial Market News”. Esse preditor será denotado por W_t . Depois de coletar as notícias, nós as classificamos mensalmente e realizamos a contagem de palavras, construindo nossa DTM(Document Term Matrix) M_t , seguindo [Lima et al. \(2018b\)](#). Antes fizemos todo o pré processamento baseado em [Hansen et al. \(2017\)](#), a fim de identificar as colocações, pois termos como “financial market”, “stock market”, “bull market” correspondem a um único conceito econômico, apesar de estarem em palavras separadas. Também aplicamos o filtro TF-IDF, a fim de excluir tanto palavras raras bem como palavras muito frequentes. Mesmo assim, continuamos com um problema de alta dimensão, ou seja, $p \gg T$.

O procedimento de previsão é baseado na janela de estimação recursiva ([Rapach and Strauss, 2010](#)) e [Lima and Meng \(2017\)](#). Nossa janela de estimação começa com 228 observações, de janeiro de 1980 a dezembro de 1998 e se expande periodicamente à medida que avançamos. As previsões fora da amostra variam de janeiro de 1999 a dezembro de 2017, correspondendo a 228 observações. Além de todo o período fora da amostra, testamos a robustez de nossas descobertas considerando o seguinte sub período de amostra mais recente: janeiro de 2008 a dezembro de 2017. Também dividimos a

³Os dados brutos vêm da página da web de Amit Goyal (<http://www.hec.unil.ch/agoyal/>).

amostra em sub períodos de recessão e expansão, onde as datas de recessão são identificadas de acordo com a NBER⁴. Para a nosso período fora da amostra, as recessões aconteceram entre março e novembro de 2001 e entre dezembro de 2007 e junho de 2009.

3.2.2 Construindo dicionários variantes no tempo

Nesta seção, queremos construir um dicionário que varie no tempo, isto é, uma matriz $M_t^* \subset M_t$ que contenha apenas as palavras mais preditivas das notícias. O prêmio de risco é denotado por r_{t+1} . Implementamos este passo aplicando Elastic Net na seguinte equação de predição linear:

$$r_{t+1} = X_t' \beta_i + W_t' \phi_i + \epsilon_{i,t+1} \quad (3.1)$$

$\hat{\beta}_h$ e $\hat{\phi}_h$ são estimados minimizando a seguinte função objetivo:

$$\min_{\beta_i, \phi_i} \sum_t (r_t - X_{t-1}' \beta_i - W_{t-1}' \phi_i)^2 + \lambda_1 \| \phi_i \|_{\ell_1} + \lambda_2 \| \phi_i \|_{\ell_2} \quad (3.2)$$

onde W_t é um vetor de $k \times 1$ de preditores predeterminados, como preditores econômicos X_t ; W_t é um vetor de $p \times 1$ de séries temporais representando valores de contagem normalizados para as palavras/termos p listadas em M_t^5 ; $\|\cdot\|_{\ell_1}$ e $\|\cdot\|_{\ell_2}$ são as normas ℓ_1 e ℓ_2 , respectivamente. Na literatura estatística, uma combinação das normas de restrição ℓ_1 e ℓ_2 é conhecido como “Elastic Net” (Zou and Hastie, 2005). Como estamos considerando um dicionário que varia no tempo, a dimensão de W_t será muito grande, ou seja $p \gg T$, e por essa razão ϕ_h serão os únicos coeficientes penalizados em (3.2). Os valores ótimos de λ_1 e λ_2 são obtidos a partir do procedimento sugerido no pacote GLMNET R (seção de regressão linear) desenvolvido por Trevor Hastie e Junyang Qian.⁶

Nossos preditores de notícias financeiras correspondem ao(s) fator(es) comuns PC_t de W_t^* , que são calculados via componentes principais. Seleccionamos o número ideal de fatores por meio da abordagem de taxa de autovalores desenvolvida por Ahn and Horenstein (2013). Em seguida, seguimos Bai and Ng (2008) para manter apenas os fatores com p valor menor ou igual a 0.01. Por exemplo, se o número ideal é 3, então consideramos os três primeiros componentes principais, mas mantemos apenas aqueles com p valor menor ou igual a 0.01. Essa abordagem é próxima da desenvolvida por Bai

⁴As datas podem ser encontradas no seguinte site. [Datas de ciclos de negócios do The National Bureau of Economic Research \(NBER\)](#)

⁵ M_t representa a DTM como em Lima et al. (2018b)

⁶O link pode ser encontrado aqui [GLMNET](#).

and Ng (2008), que aplicou a mesma ideia a dados estruturados, ou seja, propuseram primeiro selecionar as variáveis mais preditivas de um grande conjunto de co-variáveis e, então, consideraram os fatores comuns das variáveis selecionadas como preditores. Observe que essa abordagem não impõe nenhuma restrição à porcentagem de informações “negativas” e “positivas” carregadas em cada fator. De fato, mesmo que a quantidade de informações “negativas” (“positivas”) nas notícias financeiras não se altere com o tempo, essa abordagem é capaz de selecionar diferentes porcentagens de tais informações a serem carregadas nos fatores baseados em palavras selecionadas por Elastic Net. Se as informações negativas tiverem mais poder de previsão em um determinado momento, as palavras relacionadas ao sentimento negativo serão selecionadas e os fatores serão carregados com mais informações negativas.

3.2.3 Procedimento de Previsão

Antes de explicar o procedimento de previsão, introduzimos as regressões preditivas univariadas padrão estimadas por OLS (Rapach and Strauss, 2010, Lima and Meng, 2017, Welch and Goyal, 2008). Eles são expressos como:

$$r_{t+1} = X_t' \beta_i + \epsilon_{i,t+1} \quad (3.3)$$

onde X_t é uma matriz $t \times 2$ com um preditor econômico e os 1s necessário na estimação do intercepto do modelo. Com o objetivo de testar se os resultados de Rapach and Strauss (2010) no qual a combinação de previsão dos preditores econômicos melhora a previsão, adicionamos a Combinação de Previsão(FC) dos preditores econômicos com pesos iguais, que será rotulado como FC_y . Neste modelo, combinamos os 15 modelos de preditores econômicos individuais da equação (3.3). Nosso modelo benchmark será representado pela média histórica HA(“historical average”). A média histórica será estimada através regressão onde o prêmio de risco r_t será regredido apenas contra a constante.

Para descobrir se as notícias melhoram a precisão das previsões, adicionamos a (3.3) os fatores comuns construídos a partir dos preditores de palavras W_t^* . Estes modelos de News Factors serão representados por:

$$r_{t+1} = X_t' \beta_i + PC_t' \varphi_i + \epsilon_{i,t+1} \quad (3.4)$$

onde β_i e φ_i são estimados por MQO. PC_t' representa o(s) fator(es) a partir das notícias. A fim de detectar o poder preditivo das notícias sem a influência dos preditores

econômicos, incluímos o modelo onde os preditores são apenas os fatores das notícias, dando origem ao modelo “News”, denotado por:

$$r_{t+1} = PC_t' \boldsymbol{\varphi}_i + \epsilon_{i,t+1} \quad (3.5)$$

A fim de verificar se a seleção das palavras tem importância, temos um modelo em que PC_t' é obtido a partir de W_t ao invés de W_t^* , ou seja, com todas as raízes de palavras e termos. Esse modelo será rotulado por “News factor”. Também, a fim de verificar a importância do método Elastic Net, vamos estimar a equação 3.2 por ridge e LASSO, cujos modelos serão identificados por “News ridge” e “News LASSO”. Consideramos também o modelo sum-of-the-parts(SOP) de [Ferreira and Santa-Clara \(2011\)](#), dado pela equação:

$$r_{t+1} = \overline{ge}_t^{20} + dp_t - r_{f,t+1} \quad (3.6)$$

onde \overline{ge}_t^{20} é a média móvel de 20 anos do crescimento dos lucros, e dp_t é a razão dividendo-preço e $r_{f,t+1}$ é a taxa livre de risco.

Por fim, incluímos o fator das notícias financeiras no modelo SOP, dando origem ao modelo que chamamos SOP-News. Neste modelo estimamos 3.4 com $\mathbf{X}_t' = [\mathbf{1} \quad dp_t]$ e $\beta_0 = \overline{ge}_t^{20} - r_{f,t+1}$ e $\beta_1 = 1$. O vetor de parâmetros $\boldsymbol{\varphi}_i$ não tem restrição. A restrição é necessária para que o modelo Sum-of-the-Parts seja aninhado no modelo Sum-of-the-Parts news, desta forma, usamos o pacote [lmtest](#) e a função `rls` para estimar a equação de previsão do modelo SOP-News.

3.2.4 Avaliação da Previsão

Nosso procedimento de avaliação de previsão é baseado em [Lima and Meng \(2017\)](#). A primeira medida de avaliação é o R^2 , R_{OS}^2 , que compara a previsão de um Modelo condicional fora da amostra \hat{r}_{t+1} , com o modelo benchmark(incondicional) \bar{r}_{t+1} ([Campbell and Thompson, 2007](#)). Reportamos o valor de R_{OS}^2 em termos percentuais, $R_{OS}^2(\%) = 100 \times R_{OS}^2$. Segundo, para testar a hipótese nula $R_{OS}^2 \leq 0$, aplicamos os testes [Diebold and Mariano \(2002\)](#) e [Clark and West \(2007\)](#)⁷. [Campbell and Thompson \(2007\)](#) definem o R_{OS}^2 como:

⁷As estatísticas [Diebold and Mariano \(2002\)](#) e [West \(1996\)](#) geralmente são usadas para testar a hipótese nula, $R_{OS}^2 \leq 0$ entre os modelos não aninhados fora da amostra. Para modelos aninhados, como os apresentados neste artigo, [Clark and McCracken \(2001\)](#) e [McCracken \(2007\)](#) mostram que essas estatísticas têm uma distribuição não padrão. Assim, os testes [Diebold and Mariano \(2002\)](#)(DM) e [West \(1996\)](#) podem ser severamente subdimensionados sob a hipótese nula e ter baixo poder sob a hipótese alternativa

$$R_{OS}^2 = 1 - \frac{\sum_{t=1}^{T^*} (r_t - \hat{r}_t)}{\sum_{t=1}^{T^*} (r_t - \bar{r}_t)} \quad (3.7)$$

Onde T^* é o número de previsões fora da amostra. Na literatura de previsão de retorno de ações e prêmio de risco, os valores percentuais de R_{OS}^2 fora da amostra são tipicamente pequenos, mas isto não significa que seus valores econômicos sejam insignificantes. De fato, como argumentado por [Campbell and Thompson \(2007\)](#), mesmo um R^2 muito pequeno, como 0.5% para dados mensais ou 1% para dados trimestrais, ainda pode sinalizar um retorno econômico significativo dado o aumento da previsibilidade do prêmio de risco, em termos de aumento do retorno anual do portfólio para um investidor de média variância. Para estimar os ganhos econômicos das previsões, calculamos o equivalente de certeza(ou ganho de utilidade), que pode ser interpretado como a taxa de administração que um investidor estaria disposto a pagar para ter acesso às informações adicionais fornecidas pelos modelos de previsão condicional, em relação à informação disponível no modelo benchmark. Considerando um investidor avesso ao risco que tenha uma função de utilidade de média-variância e considera como otimizar a alocação da riqueza total entre um ativo de risco e um ativo livre de risco no tempo t baseado na taxa livre de risco atual r_{t+1}^f e a previsão um passo a frente do prêmio de risco, $\hat{r}_{t,t+1}$. Assim, a previsão um passo a frente do retorno é $\hat{R}_{t,t+1} = \hat{r}_{t,t+1} + \hat{r}_{t,t+1}^f$. O peso atribuído ao ativo de risco é calculado como $w = \frac{1}{\gamma} \frac{\hat{R}_{t,t+1}}{\hat{\sigma}_{t+1}^2}$, onde γ é o parâmetro de aversão ao risco e $\hat{\sigma}_{t+1}^2$ é a variância estimada do retorno. Nós impomos a restrição $w_t \in (0, 1.5)$ ⁸ para garantir que não haja venda a descoberto ([Rapach and Strauss, 2010](#)) e ([Lima and Meng, 2017](#)).

Ainda seguindo [Lima and Meng \(2017\)](#), o retorno realizado da carteira no momento $t + 1$ é $R_{t+1}^p = w_{t+1}R_{t+1} + (1 - w_{t+1})r_{t+1}^f$. sobre o período T^* fora da amostra, a utilidade de um investidor dessa alocação de carteira pode ser calculada como:

$$U = \hat{\mu}_p - \frac{1}{2}\gamma\hat{\sigma}_p^2 \quad (3.8)$$

onde $\hat{\mu}_p = \frac{1}{T^*} \sum_t R_t^p$ e $\sigma_p^2 = Var(R_t^p) = \frac{1}{T^*} \sum_t (R_t^p - \hat{\mu}_p)^2$. O ganho de utilidade é a diferença entre a utilizada baseada na previsão condicional e a utilidade da média histórica. Para facilitar a interpretação, multiplicamos os ganhos de utilidade por 1200, o que nos dá a taxa de administração anual que um investidor estaria disposto a pagar para obter acesso às informações adicionais desse modelo de previsão condicional. Essa

⁸Os pesos do ativo de risco variam de 0, onde o investidor comprará apenas ativos livres de risco a 1.5, onde o investidor assume uma posição alavancada. Se o peso for igual a 1, o investidor comprará apenas ativos de risco, e não investirá no ativo livre de risco. Se o peso for maior que 1, significa que o investidor tomará recursos emprestados a taxa livre de risco e com esse dinheiro emprestado comprará o ativo de risco, indicando um operação alavancada. Na nossa simulação não é permitido vender á descoberto.

mesma abordagem também foi usada por [Campbell and Thompson \(2007\)](#), [Rapach and Strauss \(2010\)](#), [Lima and Meng \(2017\)](#) entre outros.

Também decomparamos o erro quadrático médio de previsão ($MSFE$) em duas partes: a variância da previsão e o viés quadrado. Calculamos o $MSFE$ de qualquer previsão \hat{r}_{t+1} como $\frac{1}{T^*} \sum_t (r_{t+1} - \hat{r}_{t+1})^2$ e a variância da previsão incondicional como $\frac{1}{T^*} \sum_t (r_{t+1} - \frac{1}{T^*} \sum_t \hat{r}_{t+1})^2$, onde T^* é o número total de previsões fora da amostra. O viés ao quadrado da previsão é então calculado como a diferença entre $MSFE$ e variância de previsão ([Elliott and Timmermann, 2013b](#), [Rapach and Strauss, 2010](#)). A variância de previsão relativa(viés ao quadrado) é calculada como diferença entre a variância de previsão(viés ao quadrado) do i -ésimo modelo e a variância de previsão(viés ao quadrado) do modelo HA . Consequentemente, o valor da variância da previsão relativa(viés ao quadrado) para o modelo HA ⁹ é necessariamente igual a zero.

3.3 Resultados Empíricos

3.3.1 Resultados fora da amostra

Na tabela 17 reportamos os resultados das previsões condicionais. Esta tabela reporta a estatística R_{OS}^2 e sua significância através dos p valores do teste [Clark and West \(2007\)](#) teste(CW). Ela também exibe o ganho anual de utilidade Δ (anual %) associado a cada modelo de previsão e o p valor do teste [Diebold and Mariano \(2002\)](#) (DM). Os resultados para todo o período fora da amostra 1999:1–2017:12 confirmam que poucos modelos de previsão de preditores individuais têm R_{OS}^2 positivos e significativos. Dado que o preditor SVAR apresentou a melhor performance, acrescentamos ao mesmo a informação das notícias financeiras, a fim de verificar se as notícias melhoram a previsão. Observamos que ao acrescentar as informações das notícias financeiras, ocorre uma melhora na previsão, dado que o modelo SVAR News tem um R_{OS}^2 de 7.31%, maior que o do modelo apenas com a variável SVAR.

Ainda analisando a tabela 17 verificamos que tanto a seleção dos termos/raízes de palavras como o método de estimação importam. Veja que o modelo News Factor, em que a construção dos fatores é feita com base em todos os termos/raízes de palavras, ou seja em W_t ao invés de W_t^* não é melhor que nenhum modelo que seleciona as palavras. Assim, a seleção das palavras é fundamental na melhora da previsão. Com relação a importância do método, verificamos que o Elastic Net tem uma melhor performance que o Ridge News e o LASSO News, dado que Ridge e LASSO são aninhados no Elastic Net. Na classe de outros modelos, temos o modelo sum-of-the-parts baseado em [Ferreira](#)

⁹O modelo benchmark é a média histórica HA(historical average).

and Santa-Clara (2011). Os nossos resultados estão em linha com Rapach and Zhou (2013) que encontra uma boa performance no modelo, dado que o mesmo supera a média histórica. O período fora da amostra de Rapach and Zhou (2013) compreende 1957:1 a 2010:12, ao passo que o nosso vai de 1999:1 até 2017:12. Também observamos que o modelo Sum-of-the-Parts tem um melhor desempenho ao incluirmos as informações das notícias financeiras. O modelo de combinação de previsão de preditores individuais é superado pela média histórica. Segundo Rapach and Strauss (2010) a combinação de previsão dos preditores individuais supera a média histórica para dados trimestrais, no entanto, de acordo com os resultados da presente pesquisa o mesmo não ocorre para dados mensais durante o período fora da amostra de 1999:1 a 2017:12.

A tabela 17 também reporta os resultados do período fora da amostra entre 2008:1 a 2017:12. Esse período inicia-se com crise do sub prime a partir de 2008 e até 2009, e logo após um longo mercado de alta (bull market) que inicia-se em meados de 2009 e mantém-se até 2017¹⁰. Para esse período fora da amostra observamos que os preditores individuais em sua maioria não apresentam um bom desempenho, alguns como INFL e DE chegam a ter um desempenho pior do que em todo o período fora da amostra. O preditor SVAR é uma das exceções e tem uma melhora na performance. Com relação a classe de News Models, modelos que incluem o fator nas notícias financeiras, notamos que todos tem um R_{OS}^2 maior do que em todo o período fora da amostra, o que poder ter ocorrido devido ao aumento do poder de previsão das notícias em períodos de crise. A tabela 18 reportará os resultados para os períodos de expansão e recessão e nos ajudará a confirmar ou não essa hipótese.

Como dito anteriormente, a tabela 18 reporta os resultados para os sub períodos de recessão e expansão. As datas sobre ciclo de negócios para identificar os períodos de recessão e expansão foram obtidos na NBER. Para o sub período de expansão apenas o preditor SVAR performa melhor que a média histórica. Notamos também que os modelos que incluem a informação das notícias financeiras News e News LASSO superam a média histórica, dado que tem um R_{OS}^2 positivo. Para o período de recessão, alguns preditores individuais com DY superam a média histórica, enquanto preditores como INFL por exemplo pioram seu desempenho na recessão. Todos os preditores do grupo News Models superam a média histórica na recessão, indicando que as notícias tem um poder de previsão não só em períodos de expansão, mas também são de grande utilidade na recessão. Esses resultados confirmam os resultados de Garcia (2013) que afirma que o poder preditivo as notícias financeiras é importante especialmente em épocas de crise. O Sum-of-The-Parts é superado pela média histórica na recessão mas, ao acrescentarmos a informação das notícias, o modelo Sum-of-the-Parts News passa a superar

¹⁰Durante a escrita do artigo, em setembro de 2018 o mercado ainda é considerado de alta

a média histórica. O modelo FC é superado pelo benchmark tanto na expansão quanto na recessão.

Ainda analisando a tabela 18 verificamos que o modelo SVAR-News é o que tem a melhor performance durante os períodos de recessão e o modelo SOP-News durante os períodos de expansão. Portanto, com base nestes resultados, recomendamos ao investidor o uso do modelo SOP-News durante a expansão e o modelo SVAR-News durante recessões. Desta maneira, o ganho financeiro consegue ser maximizado. Além do mais, pode-se usar a combinação da previsão entre os dois modelos.

3.3.2 Explicando os benefícios dos modelos baseados em notícias financeiras

A figura 20 mostra a variância de previsão relativa e o viés ao quadrado relativo dos modelos de previsão, para o período fora da amostra 1999:1–2017:12. A variância da previsão relativa (viés ao quadrado) é calculada como a diferença entre a variância da previsão (viés) do i -ésimo modelo e a variância da previsão (viés ao quadrado) da média histórica (HA). Assim, o valor da variância da previsão relativa (viés ao quadrado) para o HA é necessariamente igual a zero. Cada ponto na linha representa uma previsão com o mesmo MSPE de HA; a direita da linha são previsões superadas pelo HA, e à esquerda representam previsões que superam o HA. Por fim, tanto a variância da previsão quanto o viés ao quadrado são medidos na mesma escala, então é possível determinar o trade-off entre variância e viés de cada modelo de previsão.

A média histórica é o modelo que apresenta a menor variância, mas tende a ser tendenciosa. Portanto um modelo de previsão que apresenta uma boa performance deve ser capaz de reduzir o viés mantendo a variância sob controle. O modelo SVAR News apresenta uma alta variância, mas a mesma é compensada por uma alta redução no viés. A alta variância do modelo pode ser explicado pelo maior número de preditores. Além do preditor SVAR, é incorporado ao mesmo fator(es) construídos a partir das notícias financeiras. O modelo News, apresenta variância e viés menor que o SVAR, dado que o mesmo não inclui o preditor SVAR, apenas PC_t . A principal fonte de melhora da previsão do SOP deve-se a diminuição da variância. Esse modelo tem uma variância um pouco maior que a média histórica, mas, consegue uma redução significativa no viés. Isso pode ser explicado pela restrição nos parâmetros, o que diminui a incerteza da estimação. Notamos também que ao incluirmos PC_t em SOP aumentamos a variância, mas em troca conseguimos uma redução significativa do viés. Com relação aos modelos que incluem os preditores econômicos individuais verificamos que os mesmos tem uma variância maior

do que a média histórica, mas não são capazes de reduzir o viés, o que justifica a baixa performance dos mesmos.

A Figura 21 mostra a evolução ao longo do tempo de portfólios cuja alocação foi definida com base nas previsões geradas pelos modelos HA, FC, SOP-News e News. Supomos que o valor inicial do portfólio em 1999:1 é igual a $V_t = US\$100.00$. Escolhemos do modelo SOP-News e não o SVAR-News porque apesar do modelo SVAR-News ter uma maior R_2^{OS} , ele apresenta um ganho de utilidade menor que o SOP-News. Usamos a mesma seleção de peso de Lima and Meng (2017) para calcular o Ganho de Utilidade, definindo o parâmetro de aversão ao risco $\gamma = 3$ e para $\hat{\sigma}_{t+1}$ a variância da amostra computada sobre uma janela móvel de cinco anos dos retornos históricos. Por exemplo, com base na previsão do modelo gerada para $(t + 1)$, obtemos os pesos para o ativo de risco $\hat{r}_{t,t+1}$ e para o ativo livre de risco $\hat{r}_{t,t+1}^f$. Em seguida calculamos o retorno do portfólio R_{t+1}^p , definido anteriormente na seção 3.2.4. O passo seguinte, é calcular qual será o valor do portfólio ao final do período $(t + 1)$, definido como $V_{t+1} = V_t(1 + R_{t+1}^p)$. Repetimos esse procedimento recursivamente até o final das observações fora da amostra T^* , para, assim obtermos o valor do portfólio em 2017:12.

Um portfólio escolhido com base no modelo News vale em 2017:12 em torno de US\$ 140.00 enquanto a carteira baseada em HA vale US\$ 72,44. Para o modelo de combinação de previsão um portfólio de US\$ 100.00 em 1999 valeria em 2017:12 em torno de US\$ 80.00. O melhor retorno financeiro provém do modelo SOP-News, onde um investimento inicial de US\$ proporcionou um retorno financeiro de mais de US\$ 300.00 ao final 2017¹¹. Portanto, temos evidências que o uso de notícias financeiras trazem um aumento significativo no retorno do investimento no período pesquisado. A análise da figura 21 complementa os resultados do ganho de utilidade reportados na tabela 17 e 18. Por exemplo, o investidor que baseou-se no modelo SVAR news teve em média um retorno 3.51% maior do que o retorno do portfólio baseado na média histórica anualmente. Portanto, esse resultados indicam que as notícias financeiras importam, e proporcionam um significativo ganho econômico/financeiro de longo prazo para os investidores. Por fim, destacamos que o modelo SOP news foi o que apresentou o maior ganho de utilidade ao investidor, superando a média histórica em 7% ao ano.

A tabela 19 reporta os termos/raízes de palavras com maior poder preditivo. Classificamos as palavras de acordo com o valor absoluto do coeficiente. Assim a tabela reporta o percentual em que o termo apareceu entre o top-10. Termos como “world economy” figura entre o top-10 durante todo o período fora da amostra. Também observamos termos como “banking system”, “junk bond” e “retail sale” entre outros que apareceram no top 10 com maior frequência relativa no sub período após 2008, o que

¹¹Em nossa simulação não consideramos custo de corretagem

pode indicar uma maior importância desse vocabulário após a crise do sub prime. Isso serve para ilustrar que o poder preditivo das palavras muda com o tempo, dado que pode depender do vocabulário usado ou até do surgimento de novas palavras no dicionário.

3.4 Considerações Finais

O artigo propôs verificar se as notícias financeiras adicionam informação aos modelos de previsão já conhecidos na literatura. Nosso método usa as notícias financeiras do Wall Street Journal e do The New York Times. Também utilizamos técnicas de Aprendizado de Máquina Supervisionado como Elastic Net a fim de selecionar as palavras mais preditivas. Nossos resultados confirmam os resultados de [Rapach and Strauss \(2010\)](#), [Lima and Meng \(2017\)](#), [Rapach and Zhou \(2013\)](#) de que os preditores econômicos individuais não acrescentam informação a média histórica. No entanto, quando acrescentamos as informações dos fatores de notícias financeiras, a modelos com preditores como SVAR ou SOP observamos uma melhora significativa. Nossos resultados também estão em linha com [Rapach and Zhou \(2013\)](#), que encontram que o modelo Sum-of-the-parts supera a média histórica. No entanto, mostramos que ao acrescentarmos a esse modelo o(s) fator(es) provenientes das notícias financeiras, seu desempenho é potencializado. Também calculamos um modelo apenas com a informação das notícias e descobrimos que o mesmo supera a média histórica por uma margem substancial.

Outra inovação do presente artigo foi a utilização de um dicionário variante no tempo ao invés de um dicionário fixo. Portanto, concluímos que o uso de técnicas de aprendizado de máquina supervisionado com preditores baseados em notícias financeiras traz ganhos estatísticos e econômicos a previsão condicional. Para futuras pesquisas sugere-se o uso de dados de alta frequência e/ou dados diários a fim de verificar se as notícias tem o mesmo poder preditivo encontrado para dados mensais. Também pode-se efetuar o estudo se o poder preditivo encontrado para o índice de ações como o S&P 500 se mantém em nível de firmas individuais.

Conclusão

A tese objetivou encontrar novos métodos e preditores que melhorassem a acurácia da previsão fora da amostra de variáveis macroeconômicas como o PIB e financeiras como o prêmio de risco do mercado. Objetivamos aliar dados não estruturados com métodos de aprendizado de máquina aos preditores econômicos tradicionais. O uso dessa nova categoria de dados vem sendo largamente utilizado por modelos de previsão econômica, devido à rapidez com que se pode obter a informação.

Com relação ao primeiro capítulo, encontramos que o índice de incerteza construído para a economia brasileira pode ser usado como um indicador antecedente a fim de prever recessões, pois demonstramos que o mesmo tem uma correlação negativa com a produção industrial menor que a volatilidade do mercado de ações. Neste sentido, a principal contribuição do primeiro capítulo da tese aos formuladores de políticas, por exemplo, é a indicação de um novo índice para a economia brasileira que antecipa futuras recessões econômicas. De posse do índice, o gestor pode adotar uma política econômica contracíclica, antecipando reduções da taxa de juros, por exemplo.

No segundo capítulo descobrimos que as palavras importam e tem um poder de previsão não trivial. Ao usarmos aprendizado de máquina supervisionado para selecionarmos as palavras e termos mais preditivos, obtemos uma melhora significativa na previsão da taxa de crescimento do Produto dos Estados Unidos. A principal contribuição do segundo capítulo no que diz respeito a formulação de políticas é que a previsão macroeconômica pode ser melhorada ao incluirmos as informações das minutas do FED. De posse de uma previsão mais acurada, os formuladores de política podem tomar uma melhor decisão, trazendo assim, um aumento do bem estar da sociedade. As previsões também podem ajudar a investidores e empresários, que dependem da previsão do Produto para tomar decisões de investimento.

Por fim, o terceiro capítulo evidencia que o uso de notícias em modelos de previsão traz ganhos econômicos-financeiros significativos, superando modelos tradicionalmente usados na literatura. Nesse sentido, a tese contribui tanto para investidores individuais quanto para grandes gestores de carteira em melhorar a previsão do retorno e, assim,

tomar uma melhor decisão, dado que nosso modelo tem uma boa performance tanto na expansão quanto na recessão. Assim, nosso modelo tanto performa bem ao passar a mensagem ao investidor que em dado mês ou meses a melhor decisão é ficar fora de ativos de risco, como também no mercado de alta o modelo indica que o investidor deve alavancar-se a fim de maximizar o retorno.

Para futuras pesquisas sugere-se por exemplo, adicionar ao preditores baseados em palavras as variáveis de alta frequência para assim, verificar se a previsão do produto torna-se melhor ainda. Também outra pesquisa interessante seria verificar se as atas do COMPOM no Brasil também tem o mesmo poder preditivo que as minutas do FED. Com relação ao capítulo 3 sugere-se usar uma frequência diária ou até mesmo intra diária, a fim de verificar se o poder preditivo as notícias financeiras se mantém.

Referências Bibliográficas

- Kyle Jurado, Sydney C Ludvigson, and Serena Ng. Measuring uncertainty. *The American Economic Review*, 105(3):1177–1216, 2015.
- Matthew Gentzkow, Bryan T. Kelly, and Matt Taddy. Text as data. Working Paper 23276, National Bureau of Economic Research, March 2017. URL <http://www.nber.org/papers/w23276>.
- Diego Garcia. Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300, 2013.
- Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *Journal of computational science*, 2(1):1–8, 2011.
- Rüdiger Bachmann, Steffen Elstner, and Eric R Sims. Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, 5(2):217–49, 2013.
- João Victor Issler, Hilton Hostalacio Notini, and Claudia Fontoura Rodrigues. Constructing coincident and leading indices of economic activity for the brazilian economy. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2012(2):43–65, 2013.
- Nicholas Bloom. The impact of uncertainty shocks. *econometrica*, 77(3):623–685, 2009.
- Andrea Carriero, Todd E Clark, and Massimiliano Marcellino. Common drifting volatility in large bayesian vars. *Journal of Business & Economic Statistics*, 34(3):375–390, 2016.
- Chiara Scotti. Surprise and uncertainty indexes: Real-time aggregation of real-activity macro-surprises. *Journal of Monetary Economics*, 82:1–19, 2016.
- Nicholas Bloom, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J Terry. Really uncertain business cycles. Technical report, National Bureau of Economic Research, 2012.
- Eric C So. A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts? *Journal of Financial Economics*, 108(3):615–640, 2013.

- Thomas J Sargent, Christopher A Sims, et al. Business cycle modeling without pretending to have too much a priori economic theory. *New methods in business cycle research*, 1:145–168, 1977.
- Robert F Engle and Tim Bollerslev. Modelling the persistence of conditional variances. *Econometric reviews*, 5(1):1–50, 1986.
- Wagner Piazza Gaglianone, Luiz Renato Lima, Oliver Linton, and Daniel R Smith. Evaluating value-at-risk models via quantile regression. *Journal of Business & Economic Statistics*, 2012.
- Francis X Diebold and Glenn D Rudebusch. A nonparametric investigation of duration dependence in the american business cycle. *Journal of Political Economy*, pages 596–616, 1990.
- Joseph E Engelberg and Christopher A Parsons. The causal impact of media in financial markets. *The Journal of Finance*, 66(1):67–97, 2011.
- Asad Dossani. *Essays on Inference from Option Markets*. PhD thesis, UC San Diego, 2018.
- Michelle T Armesto, RUBÉN HERNÁNDEZ-MURILLO, Michael T Owyang, and Jeremy Piger. Measuring the information content of the beige book: A mixed data sampling approach. *Journal of Money, Credit and Banking*, 41(1):35–55, 2009.
- Ellyn Boukus and Joshua V Rosenberg. The information content of fomc minutes. *New York Fed*, 2006.
- Stephen G Cecchetti et al. What the fomc says and does when the stock market booms. In *Asset Prices and Monetary Policy, Proceedings of the Research Conference of the Reserve Bank of Australia*, pages 77–96, 2003.
- Jonathan H Wright. What does monetary policy do to long-term interest rates at the zero lower bound? *The Economic Journal*, 122(564), 2012.
- Carlo Altavilla and Domenico Giannone. The effectiveness of non-standard monetary policy measures: Evidence from survey data. *Journal of Applied Econometrics*, 32(5): 952–964, 2017.
- Leif Anders Thorsrud. Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, pages 1–17, 2018.
- Stephen Hansen, Michael McMahon, and Andrea Prat. Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870, 2017.

- Chiranjit Chakraborty and Andreas Joseph. Machine learning at central banks. *Bank of England*, 2017.
- Marcelle Chauvet and Simon Potter. Forecasting output. In *Handbook of Economic Forecasting*, volume 2, pages 141–194. Elsevier, 2013.
- Jushan Bai and Serena Ng. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317, 2008.
- Jiahua Li, Ilias Tsiakas, and Wei Wang. Predicting exchange rates out of sample: Can economic fundamentals beat the random walk? *Journal of Financial Econometrics*, 13(2):293–341, 2015.
- Luiz Renato Lima, Fanning Meng, and Lucas Lúcio Godeiro. Quantile forecasting with mixed-frequency data. *International Journal of Forecasting*, Forthcoming, 2018a.
- Lawrence J Christiano, Martin Eichenbaum, and Charles L Evans. Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics*, 1: 65–148, 1999.
- Valerie A Ramey. Macroeconomic shocks and their propagation. In *Handbook of Macroeconomics*, volume 2, pages 71–162. Elsevier, 2016.
- Ralf Ahrens. Predicting recessions with interest rate spreads: a multicountry regime-switching analysis. *Journal of international Money and Finance*, 21(4):519–537, 2002.
- Andrew Ang, Monika Piazzesi, and Min Wei. What does the yield curve tell us about gdp growth? *Journal of econometrics*, 131(1-2):359–403, 2006.
- Henri Bernard and Stefan Gerlach. Does the term structure predict recessions? the international evidence. *International Journal of Finance & Economics*, 3(3):195–215, 1998.
- Agustín Duarte, Ioannis A Venetis, and Ivan Paya. Predicting real growth and the probability of recession in the euro area using the yield spread. *International Journal of Forecasting*, 21(2):261–277, 2005.
- Arturo Estrella. Why does the yield curve predict output and inflation? *The Economic Journal*, 115(505):722–744, 2005.
- Arturo Estrella and Gikas A Hardouvelis. The term structure as a predictor of real economic activity. *The journal of Finance*, 46(2):555–576, 1991.
- Arturo Estrella and Frederic S Mishkin. Predicting us recessions: Financial variables as leading indicators. *Review of Economics and Statistics*, 80(1):45–61, 1998.

- Charles I Plosser and K Geert Rouwenhorst. International term structures and real economic growth. *Journal of monetary economics*, 33(1):133–155, 1994.
- James H Stock and Mark W Watson. New indexes of coincident and leading economic indicators. *NBER macroeconomics annual*, 4:351–394, 1989.
- Arman Mansoorian and Mohammed Mohsin. Monetary policy in a cash-in-advance economy: employment, capital accumulation, and the term structure of interest rates. *Canadian Journal of Economics/Revue canadienne d'économique*, 37(2):336–352, 2004.
- Finn E Kydland and Edward C Prescott. Rules rather than discretion: The inconsistency of optimal plans. *Journal of political economy*, 85(3):473–491, 1977.
- David O Lucca and Emanuel Moench. The pre-fomc announcement drift. *The Journal of Finance*, 70(1):329–371, 2015.
- Kristina Toutanova, Dan Klein, Christopher D Manning, and Yoram Singer. Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, pages 173–180. Association for Computational Linguistics, 2003.
- Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.
- Graham Elliott and Allan Timmermann. *Handbook of economic forecasting*. Elsevier, 2013a.
- Francis X Diebold and Minchul Shin. Machine learning for regularized survey forecast combination: Partially-egalitarian lasso and its derivatives. Technical report, National Bureau of Economic Research, 2018.
- Bradley Efron, Trevor Hastie, Iain Johnstone, Robert Tibshirani, et al. Least angle regression. *The Annals of statistics*, 32(2):407–499, 2004.
- Silvia Gonçalves, Michael W McCracken, and Benoit Perron. Tests of equal accuracy for nested models with estimated factors. *Journal of Econometrics*, 198(2):231–252, 2017.
- Seung C Ahn and Alex R Horenstein. Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3):1203–1227, 2013.
- Mikael Apel and Marianna Grimaldi. The information content of central bank minutes. Technical report, Sveriges Riksbank Working Paper Series no.261, 2012.

- David O Lucca and Francesco Trebbi. Measuring central bank communication: an automated approach with application to fomc statements. Technical report, National Bureau of Economic Research, 2009.
- Charles R Nelson. The prediction performance of the frb-mit-penn model of the us economy. *The American Economic Review*, 62(5):902–917, 1972.
- Todd E Clark and Kenneth D West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, 138(1):291–311, 2007.
- Francis X Diebold and Robert S Mariano. Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134–144, 2002.
- Whitney K Newey and Kenneth D West. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix, 1986.
- Elena Andreou, Eric Ghysels, and Andros Kourtellis. Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, 31(2):240–251, 2013.
- Andrea Carriero, Todd E Clark, and Massimiliano Marcellino. Realtime nowcasting with a bayesian mixed frequency model with stochastic volatility. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(4):837–862, 2015.
- Massimiliano Marcellino, Mario Porqueddu, and Fabrizio Venditti. Short-term gdp forecasting with a mixed-frequency dynamic factor model with stochastic volatility. *Journal of Business & Economic Statistics*, 34(1):118–127, 2016.
- Graham Elliott and Allan Timmermann. *Handbook of economic forecasting*. Elsevier, 2013b.
- David E Rapach and Jack K Strauss. Bagging or combining (or both)? an analysis based on forecasting us employment growth. *Econometric Reviews*, 29(5-6):511–533, 2010.
- Luiz Renato Lima and Fanning Meng. Out-of-sample return predictability: A quantile combination approach. *Journal of Applied Econometrics*, 32(4):877–895, 2017.
- Bruce E Hansen and Byeongseon Seo. Testing for two-regime threshold cointegration in vector error-correction models. *Journal of econometrics*, 110(2):293–318, 2002.
- Ivo Welch and Amit Goyal. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508, 2007.
- David Rapach and Guofu Zhou. Forecasting stock returns. In *Handbook of economic forecasting*, volume 2, pages 328–383. Elsevier, 2013.

- Miguel A Ferreira and Pedro Santa-Clara. Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, 100(3):514–537, 2011.
- Tim Loughran and Bill McDonald. The use of word lists in textual analysis. *Journal of Behavioral Finance*, 16(1):1–11, 2015.
- Luiz Renato Lima, Lucas Lúcio Godeiro, and Mohammed Mohsin. Time-varying dictionary and the predictive power of fed minutes. In *2018 CIRET biennial international conference*, 2018b.
- David E Rapach, Matthew C Ringgenberg, and Guofu Zhou. Short interest and aggregate stock returns. *Journal of Financial Economics*, 121(1):46–65, 2016.
- Joel Peress. The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *The Journal of Finance*, 69(5):2007–2043, 2014.
- Vegard Høghaug Larsen and Leif Anders Thorsrud. Asset returns, news topics and media effects. Technical report, BI Norwegian Business School, Centre for applied macro - and petroleum economics (CAMP), 2017.
- Steven L Heston and Nitish Ranjan Sinha. News versus sentiment: Predicting stock returns from news stories. Technical report, Social Science Research Network (SSRN), 2015.
- Ivo Welch and Amit Goyal. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508, 2008.
- John Y Campbell and Samuel B Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531, 2007.
- Kenneth D West. Asymptotic inference about predictive ability. *Econometrica: Journal of the Econometric Society*, pages 1067–1084, 1996.
- Todd E Clark and Michael W McCracken. Tests of equal forecast accuracy and encompassing for nested models. *Journal of econometrics*, 105(1):85–110, 2001.
- Michael W McCracken. Asymptotics for out of sample tests of granger causality. *Journal of Econometrics*, 140(2):719–752, 2007.

Apêndice A - Medindo Incerteza

Macroeconômica para o Brasil

TABELA 1: Séries Macroeconômicas Utilizadas na Construção do Índice de Incerteza

Código	Série
pap	Produção de Papelão
emp	Emprego - Núm. de pessoas com mais de 10 anos que tem trabalho
ren	Renda do trabalho
pi	Produção Industrial
BASE_R	Base Monetária
M1_R	Depósitos à vista
TTROCA	Termos de Troca
IBOV	Ibovespa
PROD_BCD	Produção de Bens de Consumo Duráveis
PROD_BC	Produção de Bens de Consumo
PROD_BCI	Produção de Bens de Consumo Intermediários
PROD_BK	Produção de Bens de Capital
CAMBIO	Taxa de Câmbio Nominal
PROD_CAM	Produção de Caminhões
PROD_ONI	Produção de Ônibus
PRODAUTO	Produção de Automóveis
PROD_VE	Produção de Veículos
HTP	Horas trabalhadas na produção
HPP	Horas pagas na produção
PO	Pessoal Ocupado na Indústria
SAL_R	Salário real da indústria
ICMS	ICMS Arrecadado
INPC	Índice Nacional de preços ao consumidor
SELIC	Taxa de juros SELIC
EXP_QUANTUM	Quantidade das Exportações

TABELA 2: Impacto da Incerteza Individual na incerteza agregada

	h=1	h=3	h=12
Pap	0.19	0.02	0.01
Pi	0.10	0.10	0.15
ProdBK	0.05	0.04	0.05
INPC	0.01	0.01	0.01
Ibov	0.02	0.33	0.01

TABELA 3: Teste de Causalidade de Granger Incerteza x Produção Industrial

	U1(2 lags)			U3(2 lags)			U12(3 lags)		
	Estatística F	P- Valor		Estatística F	P- Valor		Estatística F	P- Valor	
Produção Industrial não granger-causa Incerteza	1.67	0.19		2.54	0.08		1.00	0.39	
Incerteza não granger-causa produção industrial	2.67	0.07		4.92	0.00		3.39	0.02	

TABELA 4: Teste de Causalidade de Granger Incerteza x Emprego

	U1(2 lags)			U3(2 lags)			U12(3 lags)		
	Estatística F	P- Valor		Estatística F	P- Valor		Estatística F	P- Valor	
Emprego não granger-causa Incerteza	1.12	0.32		0.31	0.73		3.05	0.03	
Incerteza não granger-causa emprego	4.02	0.02		1.10	0.33		1.01	0.38	

TABELA 5: Decomposição da Variância para a produção industrial, emprego e horas

Lag	$\overline{\mathcal{U}(1)}$	$\overline{\mathcal{U}(3)}$	$\overline{\mathcal{U}(12)}$	Vol. Ibov
Emprego				
1	0	0	0	0
5	7.08	0.36	0.74	1.65
10	8.53	0.42	0.79	1.97
Horas				
1	0	2.07	18.18	14.56
5	3.39	1.8	14.83	11.16
10	4.65	1.93	14.93	11.53
Produção Industrial				
1	1.13	0	0	0
5	3.44	0.86	0.17	2.58
10	4.06	1.04	0.19	4.11

TABELA 6: Peso de cada série macroeconômica na incerteza agregada

Série	Pesos	Pesos Normalizados
TTROCA	0.933	0.090
EXPQUANTUM	0.891	0.086
PI	0.881	0.085
PRODONI	0.871	0.084
PROD_BCD	0.867	0.084
IBOV	0.786	0.076
PRODAUT	0.676	0.065
PRODCAM	0.662	0.064
pap	0.644	0.062
INPC	0.539	0.052
PROD_BK	0.517	0.050
PROD_BCI	0.506	0.049
ICMS	0.504	0.049
CAMBIO_R	0.466	0.045
PRODVE	0.211	0.020
SAL_R	0.193	0.019
ren	0.100	0.010
SELIC	0.081	0.008
HTT	0.077	0.007
HPP	0.074	0.007
PO	0.054	0.005
emp	0.009	0.001
PROD_BC	-0.007	-0.001
M1_R	-0.052	-0.005
BASE_R	-0.140	-0.014

TABELA 7: Teste de Causalidade para o índice de incerteza com pesos baseados em fatores

	U1(2 lags)		U3(2 lags)		U12(3 lags)	
	Estatística F	P- Valor	Estatística F	P- Valor	Estatística F	P- Valor
Produção Industrial não granger-causa Incerteza	2.74	0.07	3.23	0.04	3.59	0.03
Incerteza não granger-causa produção industrial	3.48	0.03	4.42	0.01	2.91	0.05
	U1(1 lag)		U3(1 lag)		U12(1 lag)	
Emprego não granger-causa Incerteza	Estatística F	P- Valor	Estatística F	P- Valor	Estatística F	P- Valor
Incerteza não granger-causa emprego	1.22	0.29	0.1	0.74	0.5	0.47
	4.54	0.01	3.87	0.05	3.61	0.06

TABELA 8: Turning Point Produção Industrial e Índice de Incerteza

Turning Point PI	Turning Point U1	Turning Point U3	Turning Point U12
2008.04	2008.04	2008.05	2008.09
2008.05	2008.06	2008.10	2009.10
2008.07	2008.10	2009.10	2011.06
2008.08	2009.04	2011.03	2012.07
2008.09	2009.05	2011.09	2013.10
2009.01	2009.08	2011.11	2014.10
2009.03	2010.08	2012.12	2015.05
2009.04	2013.05	2013.09	
2009.12	2013.12	2014.03	
2010.03	2014.04	2014.05	
2010.10	2014.09	2015.01	
2011.04	2015.04		
2012.05			
2012.10			
2012.12			
2013.03			
2013.04			
2013.12			
2014.02			
2015.02			

TABELA 9: QPS(1) Score

	u1	u3	u12
QPS	29.55%	30.68%	28.41%

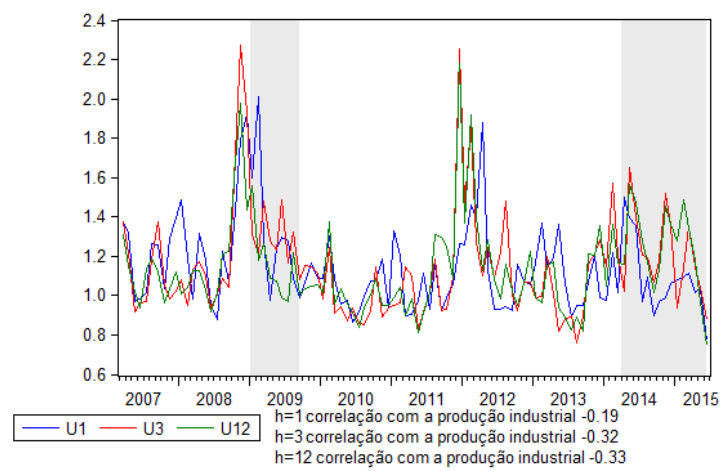


FIGURA 1: Incerteza Agregada \bar{U}_t^y para $h=1,3,12$.

O crescimento da produção industrial é computado como a média móvel de 12 meses. A área em cinza representa os períodos de recessão.

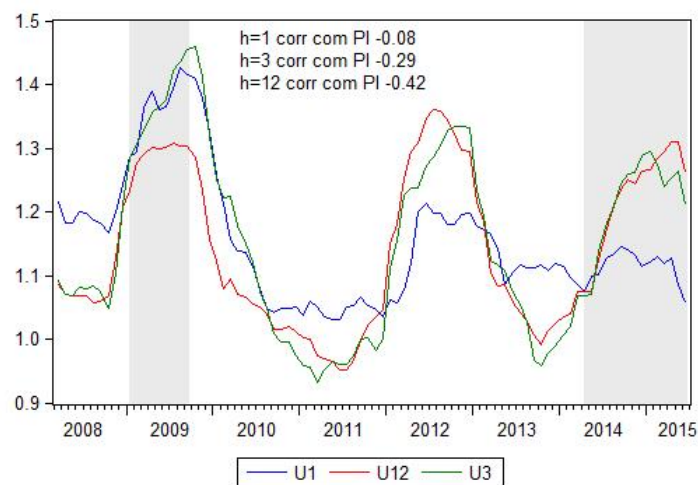


FIGURA 2: Média Móvel 12 meses para a incerteza agregada $h=1,3,12$.

As áreas em cinza representam os trimestres em recessão para a Economia Brasileira.

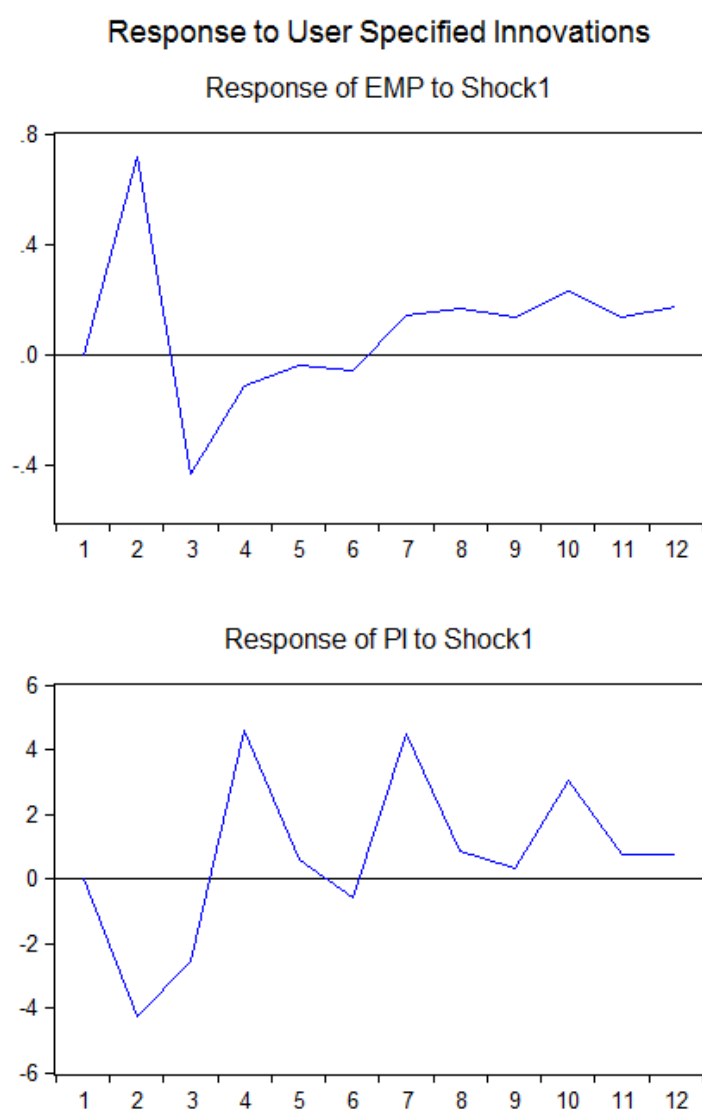


FIGURA 3: Resposta da produção industrial e do emprego a choques na volatilidade do IBOVESPA

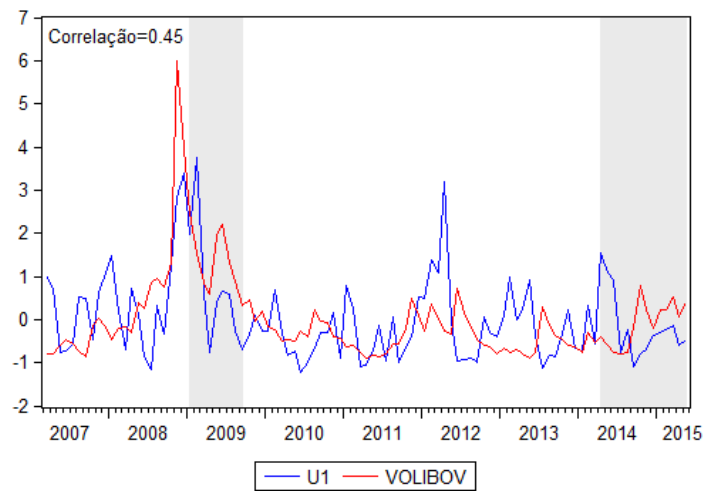


FIGURA 4: Volatilidade do IBOVESPA e Medida de Incerteza Macroeconômica. As variáveis foram padronizadas na distribuição normal.

As variáveis foram padronizadas de acordo com a distribuição normal.

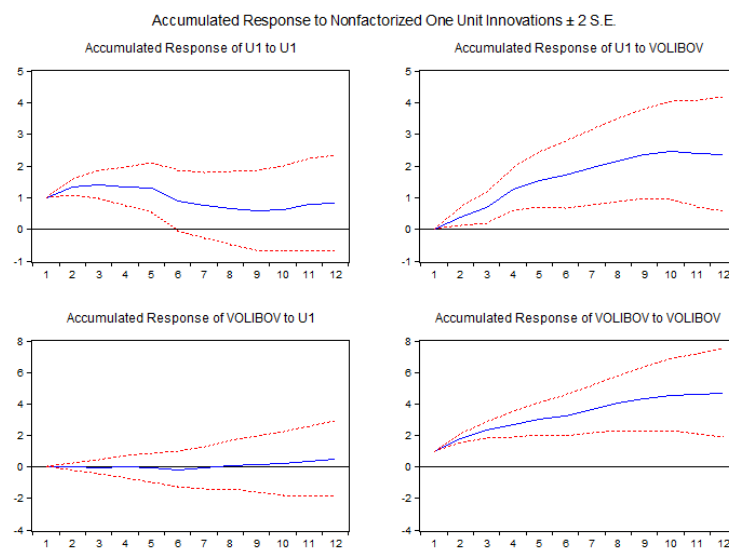


FIGURA 5: Impulso Resposta Incerteza $h=1$ e Volatilidade do IBOVESPA

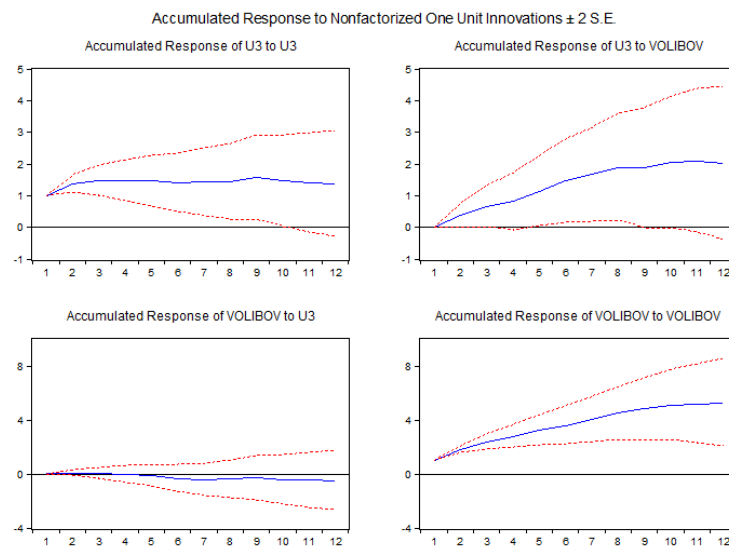


FIGURA 6: Impulso Resposta Incerteza $h=3$ e Volatilidade do IBOVESPA

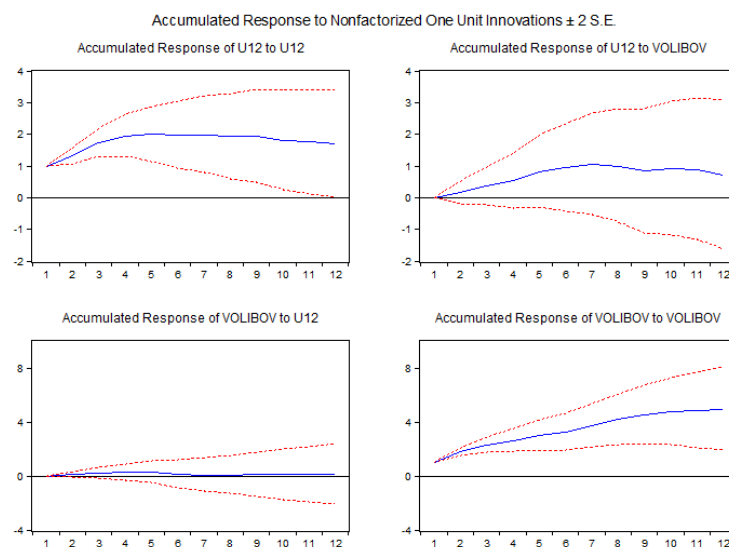


FIGURA 7: Impulso Resposta Incerteza $h=12$ e Volatilidade do IBOVESPA

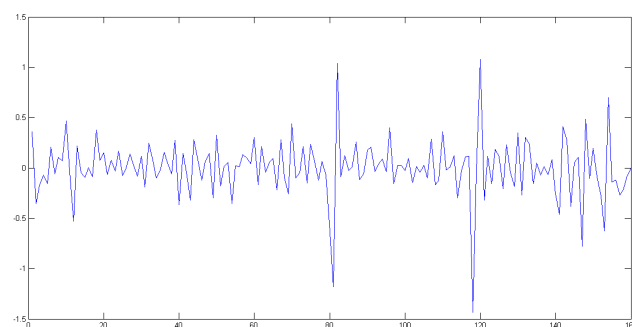


FIGURA 8: Fator Macroeconômico construído a partir do conjunto de variáveis

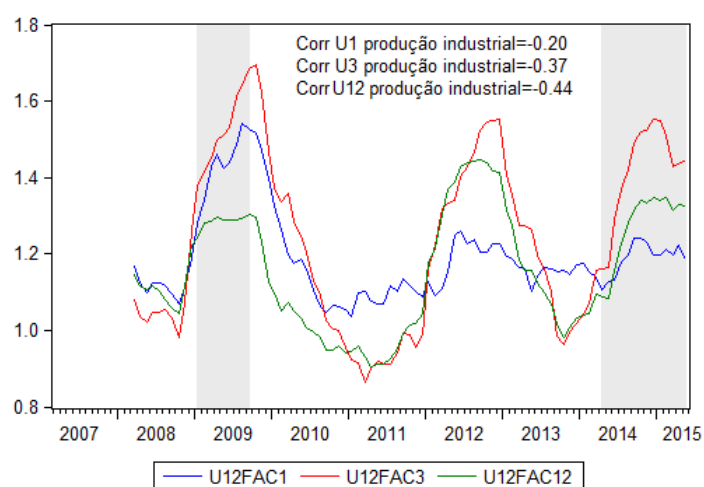


FIGURA 9: Índice de Incerteza Macroeconômica Agregado construído a partir dos pesos obtidos via análise de componentes principais.

Os valores representam a média móvel de 12 meses. A correlação foi calculada entre a média móvel do índice 12 meses e a média móvel da taxa de crescimento da produção industrial para 12 meses.

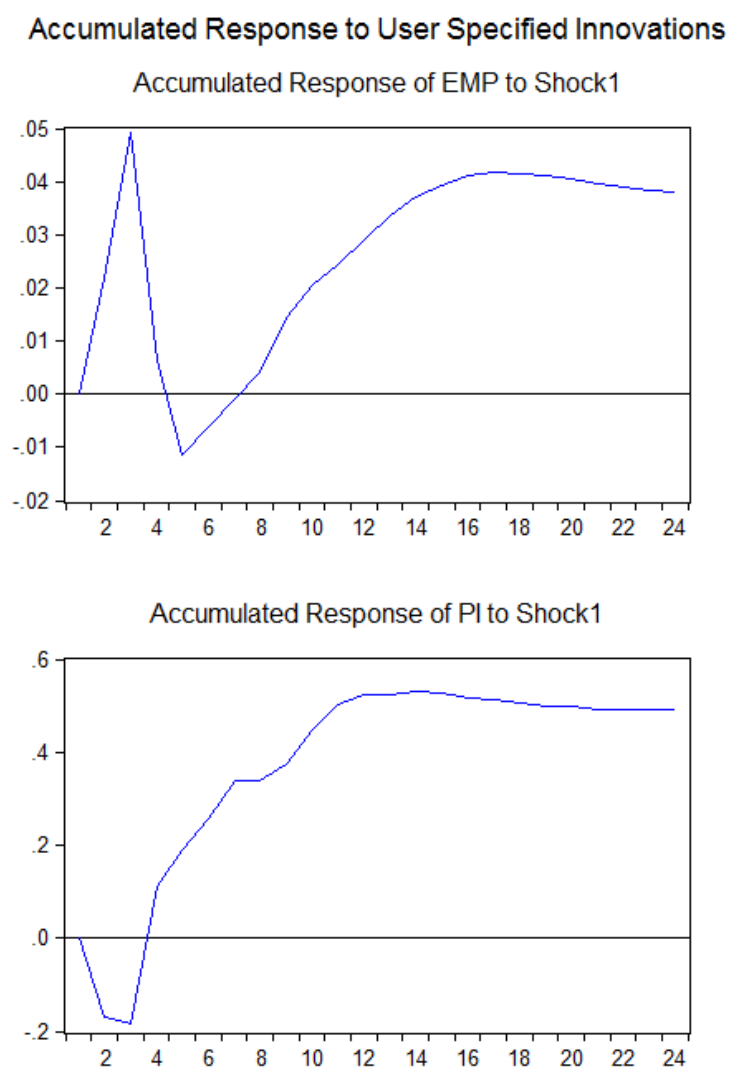


FIGURA 10: Resposta da Produção Industrial e do Emprego a um choque na Incerteza

Accumulated Response to User Specified Innovations

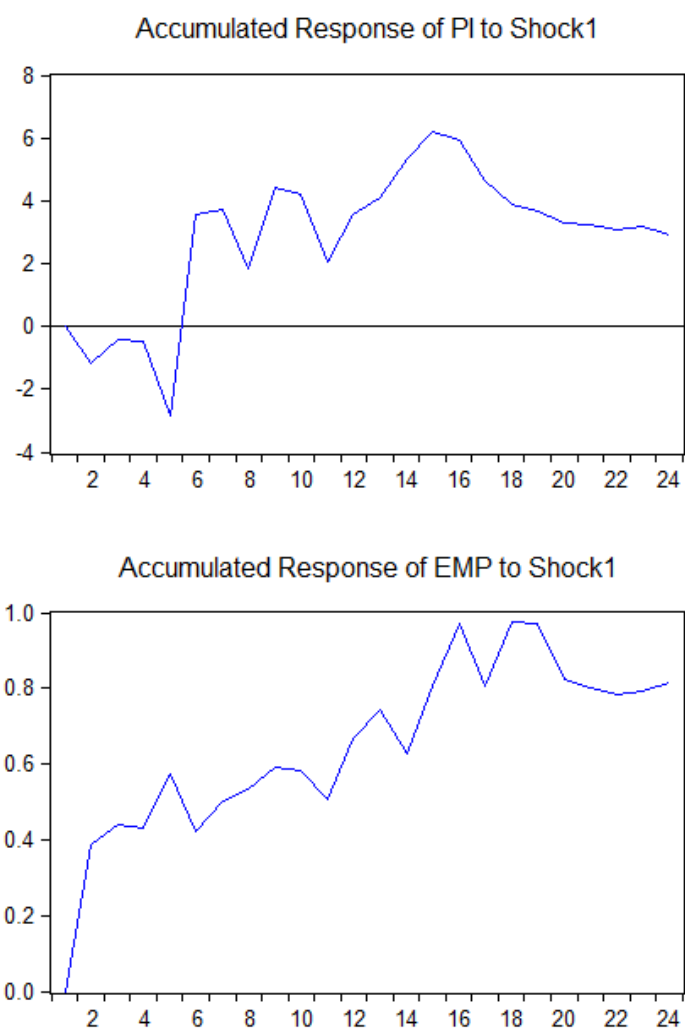


FIGURA 11: Resposta da Produção Industrial e do Emprego a um choque no Índice de Incerteza com pesos baseados nos fatores $h=3$

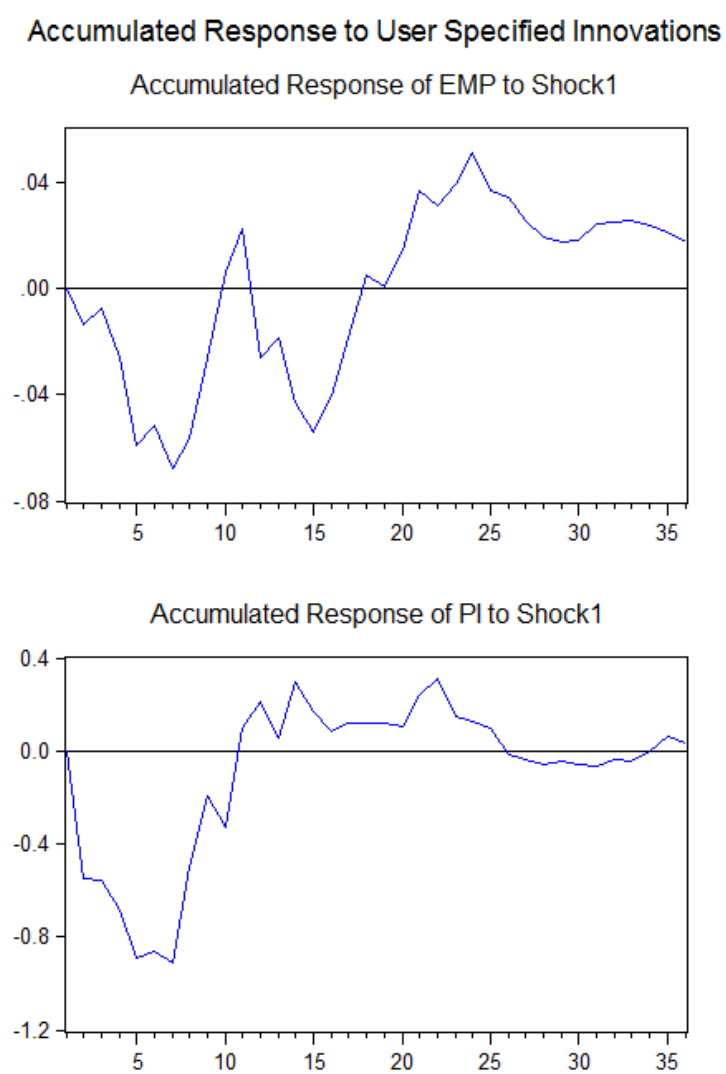


FIGURA 12: Resposta da Produção Industrial e do Emprego a um choque no Índice de Incerteza com pesos baseados nos fatores $h=12$

Apêndice B - Identificando o Poder Preditivo das Minutas do FED.

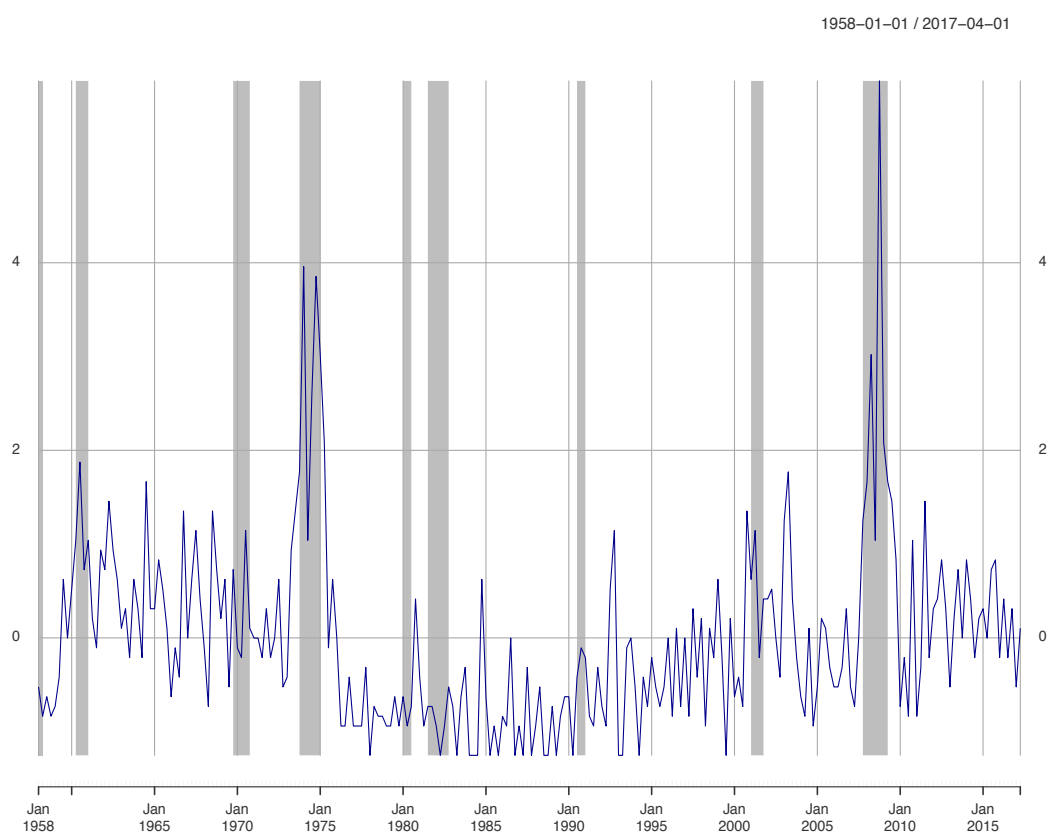


FIGURA 13: Gráfico da série de tempo “economic activity”, 1958(T1)-2017(T2).

Esta figura mostra o gráfico da série de tempo da contagem normalizada do termo “economic activity” no período de 1958(T1)-2017(T2). A área em cinza representa os períodos de recessão de acordo com a NBER.

TABELA 10: Data de Divulgação do Produto e das Minutas do FED

Trimestre	Data de Divulgação do Produto	Data de Divulgação das Minutas do FED
1992T1	29-mai-92	22-mai-92
1992T2	27-ago-92	21-ago-92
1992T3	27-out-92	9-out-92
1992T4	27-fev-93	5-fev-93
1993T1	28-mai-93	21-mai-93
1993T2	29-jul-93	12-jul-93
1993T3	27-out-93	8-out-93
1993T4	1-mar-94	7-fev-94
1994T1	27-mai-94	20-mai-94
1994T2	29-jul-94	11-jul-94
1994T3	28-out-94	13-out-94
1994T4	27-jan-95	16-jan-95
1995T1	31-mai-95	1-mai-95
1995T2	28-jul-95	10-jul-95
1995T3	27-out-95	17-out-95
1995T4	23-fev-96	5-fev-96
1996T1	30-mai-96	24-mai-96
1996T2	1-ago-96	5-jul-96
1996T3	27-nov-96	15-nov-96
1996T4	28-fev-97	7-fev-97
1997T1	26-mai-97	23-mai-97
1997T2	31-jul-97	7-jul-97
1997T3	26-nov-97	14-nov-97
1997T4	27-fev-98	6-fev-98
1998T1	27-mai-98	22-mai-98
1998T2	31-jul-98	6-jul-98
1998T3	30-out-98	19-out-98
1998T4	26-fev-99	5-fev-99
1999T1	27-mai-99	21-mai-99
1999T2	29-jul-99	15-jul-99
1999T3	29-out-99	8-out-99
1999T4	28-jan-00	13-jan-00
2000T1	27-abr-00	10-abr-00
2000T2	28-jul-00	14-jul-00
2000T3	27-out-00	6-out-00
2000T4	31-jan-01	6-jan-01

2001T1	27-abr-01	13-abr-01
2001T2	27-jul-01	12-jul-01
2001T3	31-out-01	5-out-01
2001T4	30-jan-02	11-jan-02
2002T1	26-abr-02	3-abr-02
2002T2	31-jul-02	12-jul-02
2002T3	31-out-02	10-out-02
2002T4	31-jan-03	11-jan-03
2003T1	25-abr-03	4-abr-03
2003T2	28-ago-03	15-ago-03
2003T3	30-out-03	2-out-03
2003T4	30-jan-04	5-jan-04
2004T1	26-mai-04	6-mai-04
2004T2	31-ago-04	12-ago-04
2004T3	30-nov-04	11-nov-04
2004T4	28-jan-05	4-jan-05
2005T1	28-abr-05	12-abr-05
2005T2	29-jul-05	21-jul-05
2005T3	28-out-05	11-out-05
2005T4	27-jan-06	3-jan-06
2006T1	28-abr-06	18-abr-06
2006T2	28-jul-06	20-jul-06
2006T3	27-out-06	11-out-06
2006T4	31-jan-07	3-jan-07
2007T1	27-abr-07	11-abr-07
2007T2	27-jul-07	19-jul-07
2007T3	31-out-07	9-out-07
2007T4	30-jan-08	2-jan-08
2008T1	30-abr-08	8-abr-08
2008T2	31-jul-08	16-jul-08
2008T3	30-out-08	7-out-08
2008T4	30-jan-09	6-jan-09
2009T1	29-abr-09	8-abr-09
2009T2	31-jul-09	15-jul-09
2009T3	29-out-09	14-out-09
2009T4	29-jan-10	6-jan-10
2010T1	30-abr-10	10-abr-10
2010T2	30-jul-10	14-jul-10
2010T3	29-out-10	12-out-10

2010T4	28-jan-11	4-jan-11
2011T1	28-abr-11	5-abr-11
2011T2	27-jul-11	12-jul-11
2011T3	27-out-11	12-out-11
2011T4	27-jan-12	3-jan-12
2012T1	27-abr-12	3-abr-12
2012T2	27-jul-12	11-jul-12
2012T3	26-out-12	4-out-12
2012T4	30-jan-13	3-jan-13
2013T1	26-abr-13	10-abr-13
2013T2	31-jul-13	10-jul-13
2013T3	7-nov-13	9-out-13
2013T4	30-jan-14	8-jan-14
2014T1	30-abr-14	9-abr-14
2014T2	30-jul-14	9-jul-14
2014T3	30-out-14	8-out-14
2014T4	30-jan-15	7-jan-15
2015T1	29-abr-15	8-abr-15
2015T2	30-jul-15	8-jul-15
2015T3	29-out-15	8-out-15
2015T4	29-jan-16	6-jan-16
2016T1	28-abr-16	8-abr-16
2016T2	28-jul-16	8-jul-16
2016T3	28-out-16	8-out-16
2016T4	27-jan-17	6-jan-17
2017T1	28-abr-17	5-abr-17
2017T2	28-jul-17	5-jul-17

Tabela 10 mostra a data de divulgação da primeira estimativa do produto e da ata da correspondente à última reunião do trimestre do FOMC.

TABELA 11: RMSFEs para $h = 1, 3, 6$ durante todo o período fora da amostra. Relativo ao Benchmark $AR(2)$.

Modelos	h=1	h=3	h=6
Modelo 1	0.748*	0.722**	0.707**
Modelo 2	0.837*	0.792**	0.771*
Modelo 3	0.852*	0.811**	0.802**
Modelo 4	1.033	1.005***	1.009
Modelo 5	0.994	1.020	0.988
Modelo 6	0.861**	0.883*	0.847**
Modelo 7	0.904***	0.894**	0.865*
Modelo 8	0.976*	0.962**	0.956*
Blue Chip	0.947**	1.011	0.985
$AR(2)$ Blue Chip	0.995**	0.992	0.977***
Combinação de Previsão	0.874*	0.832**	0.811*

Tabela 11 mostra o RMSFE dos modelos de previsão $f_{t+h,t}^j$ relativos ao benchmark $AR(2)$ para todo o período fora da amostra. Um valor menor que 1 indica que $f_{t+h,t}^j$ performa melhor que o $AR(2)$. (*), (**) and (***) denotam que a diferença entre o erro quadrado médio (MSE) do modelo $f_{t+h,t}^j$ e o benchmark é estatisticamente significativo as níveis de 1% , 5% and 10% , respectivamente, usando o teste [Clark and West \(2007\)](#). Os modelos de 1-3 representam o dicionário variante no tempo, os modelos 4 e 5 representam dicionários fixos e os modelos 6-8 representam outros preditores.

TABELA 12: RMSFEs para $h = 1, 3, 6$ durante os sub períodos de recessão e expansão.
Total e relativo ao Benchmark $AR(2)$

Modelos	h=1		h=3		h=6	
	Expansão	Recessão	Expansão	Recessão	Expansão	Recessão
AR2	1.825	3.474	1.938	4.369	2.007	4.431
Modelo 1	1.348	2.161	1.701	2.320	1.571	2.743
Relativo	0.739*	0.622*	0.878**	0.531*	0.783**	0.619*
Modelo 2	1.527	2.737	1.748	3.154	1.639	3.185
Relativo	0.837*	0.788**	0.902***	0.722*	0.817**	0.719**
Modelo 3	1.693	2.709	1.649	3.534	1.623	3.451
Relativo	0.928***	0.780**	0.851**	0.771*	0.809**	0.779*
Modelo 4	1.914	3.274	1.952	4.408	2.038	4.434
Relativo	1.048	0.942	1.007***	1.009	1.015	1.000
Modelo 5	1.823	3.043	1.995	4.070	1.974	4.229
Relativo	0.999*	0.876**	1.029	0.931***	0.983	0.954*
Modelo 6	1.600	3.123	1.688	4.159	1.686	4.205
Relativo	0.877**	0.899**	0.871	0.952**	0.840**	0.949
Modelo 7	1.655	2.918	1.688	4.172	1.790	4.001
Relativo	0.909**	0.840**	0.871*	0.955	0.892**	0.903**
Modelo 8	1.790	3.240	1.880	4.094	1.937	4.107
Relativo	0.980	0.932*	0.970*	0.937**	0.965*	0.926**
Blue Chip	1.727	2.619	1.982	4.191	1.985	4.171
Relativo	0.946**	0.754**	1.022	0.959**	0.989**	0.941**
AR(2) Blue Chip	1.817	3.312	1.933	4.134	1.979	4.214
Relative	0.995	0.953	0.997	0.946	0.986	0.950
Combinação de Previsão	1.655	2.674	1.614	3.372	1.565	3.416
Relativo	0.907*	0.771***	0.833*	0.772**	0.780*	0.771**

Tabela 12 mostra o RMSFE dos modelos de previsão $f_{t+h,t}^j$ relativos ao benchmark $AR(2)$ para todo o período fora da amostra. Um valor menor que 1 indica que $f_{t+h,t}^j$ performa melhor que o $AR(2)$. (*), (**) and (***) denotam que a diferença entre o erro quadrado médio (MSE) do modelo $f_{t+h,t}^j$ e o benchmark é estatisticamente significativo as níveis de 1% , 5% and 10% , respectivamente, usando o teste [Clark and West \(2007\)](#). Os modelos de 1-3 representam o dicionário variante no tempo, os modelos 4 e 5 representam dicionários fixos e os modelos 6-8 representam outros preditores.

TABELA 13: RMSFEs para $h = 0$ durante todo o período fora da amostra. Relativo ao Benchmark $AR(2)$.

Modelos	Amostra Completa
Modelo 1	0.702*
Modelo 2	0.772*
Modelo 3	0.808**
Modelo 4	1.027
Modelo 5	1.014
Modelo 6	0.814**
Modelo 7	0.866**
Modelo 8	0.989**
Blue Chip	0.921***
$AR(2)$ Blue Chip	0.863**
Combinação de Previsão	0.862*

Tabela 13 mostra o RMSFE dos modelos de previsão $f_{t+h,t}^j$ relativos ao benchmark $AR(2)$ para todo o período fora da amostra. Um valor menor que 1 indica que $f_{t+h,t}^j$ performa melhor que o $AR(2)$. (*), (**) and (***) denotam que a diferença entre o erro quadrado médio (MSE) do modelo $f_{t+h,t}^j$ e o benchmark é estatisticamente significativo as níveis de 1% , 5% and 10% , respectivamente, usando o teste [Clark and West \(2007\)](#). Os modelos de 1-3 representam o dicionário variante no tempo, os modelos 4 e 5 representam dicionários fixos e os modelos 6-8 representam outros preditores.

TABELA 14: RMSFEs para $h = 0$ durante os sub períodos de expansão e recessão.
Total and relativo ao Benchmark $AR(2)$.

Modelos	Expansão	Recessão
AR2	1.825	3.474
Modelo 1	1.297	2.199
Relativo	0.711*	0.633*
Modelo 2	1.595	2.602
Relativo	0.874**	0.749*
Modelo 3	1.646	2.678
Relativo	0.902**	0.771**
Modelo 4	1.984	3.192
Relativo	1.087	0.918**
Modelo 5	1.856	3.469
Relativo	1.017	0.998
Modelo 6	1.538	2.681
Relativo	0.843**	0.772**
Modelo 7	1.587	2.845
Relativo	0.870**	0.819**
Modelo 8	1.812	3.279
Relativo	0.992***	0.943*
Blue Chip	1.684	3.025
Relativo	0.923**	0.871**
AR(2) Blue Chip	1.658	2.852
Relativo	0.909**	0.821*
Combinação de Previsão	1.569	2.838
Relativo	0.860*	0.817*

Tabela 14 mostra o RMSFE dos modelos de previsão $f_{t+h,t}^j$ relativos ao benchmark $AR(2)$ para todo o período fora da amostra. Um valor menor que 1 indica que $f_{t+h,t}^j$ performa melhor que o $AR(2)$. (*), (**) and (***) denotam que a diferença entre o erro quadrado médio (MSE) do modelo $f_{t+h,t}^j$ e o benchmark é estatisticamente significativo as níveis de 1% , 5% and 10% , respectivamente, usando o teste [Clark and West \(2007\)](#). Os modelos de 1-3 representam o dicionário variante no tempo, os modelos 4 e 5 representam dicionários fixos e os modelos 6-8 representam outros preditores.

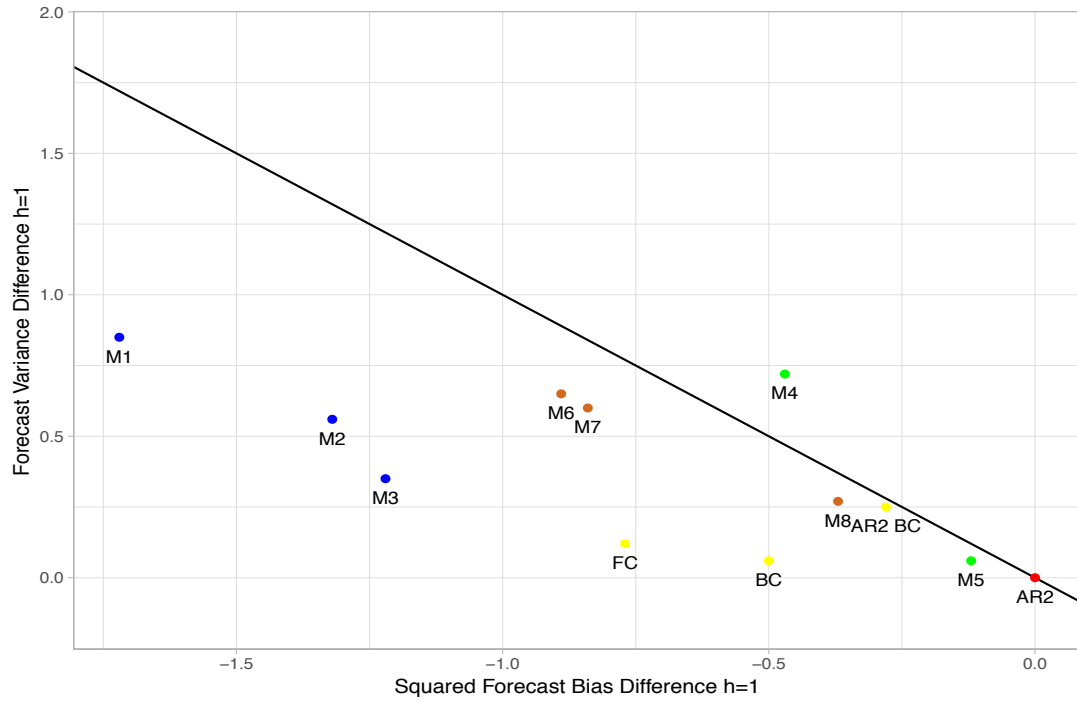


FIGURA 14: Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.

O eixo y e o eixo x representam a variância da previsão e o viés ao quadro para $h = 1$, calculados como a diferença entre a variância da previsão (viés ao quadrado) de $f_{t+h,t}^j$ a variância da previsão (viés ao quadrado) $AR(2)$. Cada ponto sobre a linha pontilhada representa uma previsão com o mesmo $MSFE$ que $AR(2)$; pontos a direita são previsões que são piores que o $AR(2)$, e pontos a esquerda são previsões melhores que o $AR(2)$. Em azul temos os modelos com o dicionários que variam no tempo; em verde temos os modelos com dicionário fixo; em amarelo temos os modelos de combinação de previsão e o Blue Chip; em laranja temos os modelos baseados em outros preditores; e em vermelho temos o Modelo benchmark $AR(2)$.

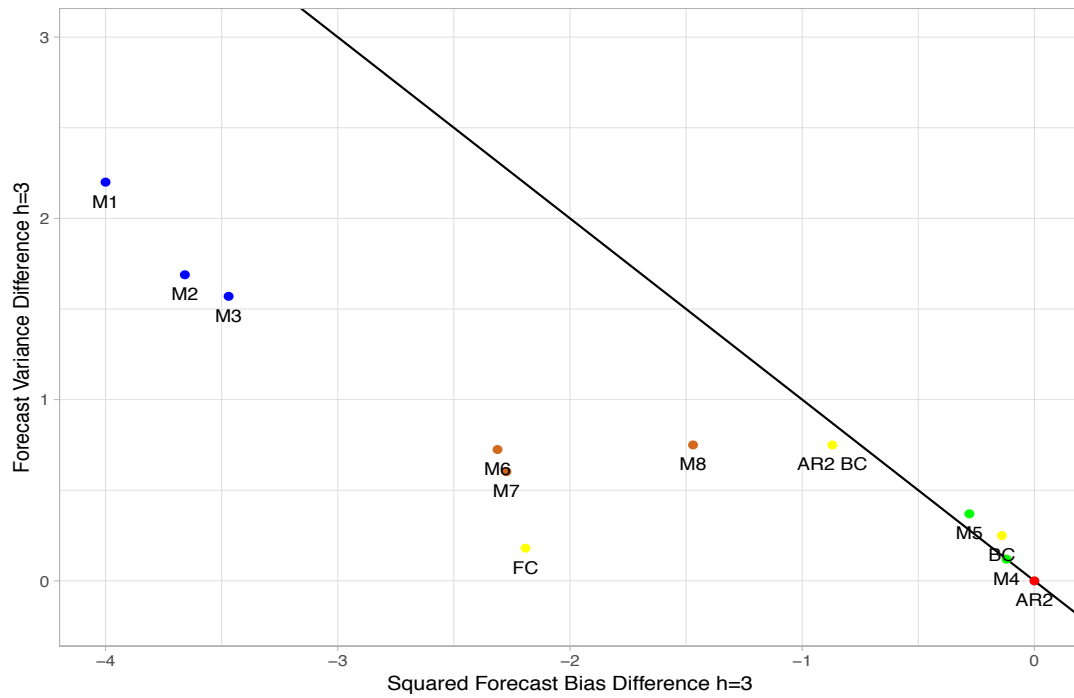


FIGURA 15: Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.

O eixo y e o eixo x representam a variância da previsão e o viés ao quadro para $h = 3$, calculados como a diferença entre a variância da previsão (viés ao quadrado) de $f_{t+h,t}^j$ a variância da previsão (viés ao quadrado) $AR(2)$. Cada ponto sobre a linha pontilhada representa uma previsão com o mesmo $MSFE$ que $AR(2)$; pontos a direita são previsões que são piores que o $AR(2)$, e pontos a esquerda são previsões melhores que o $AR(2)$. Em azul temos os modelos com o dicionários que variam no tempo; em verde temos os modelos com dicionário fixo; em amarelo temos os modelos de combinação de previsão e o Blue Chip; em laranja temos os modelos baseados em outros preditores; e em vermelho temos o Modelo benchmark $AR(2)$.

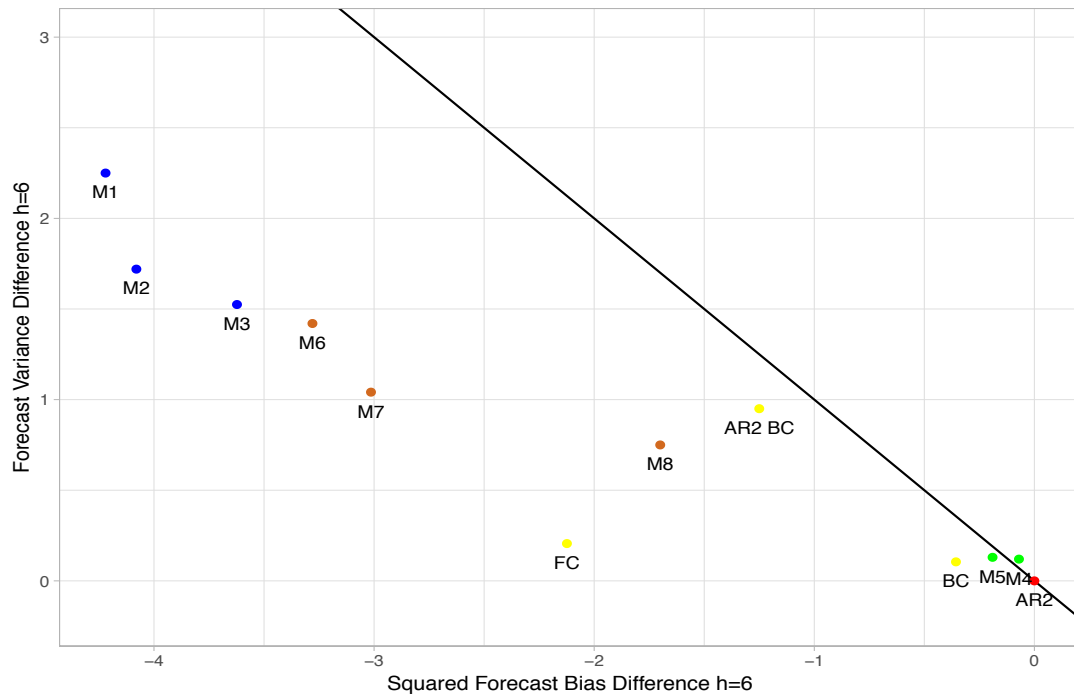


FIGURA 16: Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.

O eixo y e o eixo x representam a variância da previsão e o viés ao quadro para $h = 6$, calculados como a diferença entre a variância da previsão (viés ao quadrado) de $f_{t+h,t}^j$ a variância da previsão (viés ao quadrado) $AR(2)$. Cada ponto sobre a linha pontilhada representa uma previsão com o mesmo $MSFE$ que $AR(2)$; pontos a direita são previsões que são piores que o $AR(2)$, e pontos a esquerda são previsões melhores que o $AR(2)$. Em azul temos os modelos com o dicionários que variam no tempo; em verde temos os modelos com dicionário fixo; em amarelo temos os modelos de combinação de previsão e o Blue Chip; em laranja temos os modelos baseados em outros preditores; e em vermelho temos o Modelo benchmark $AR(2)$.

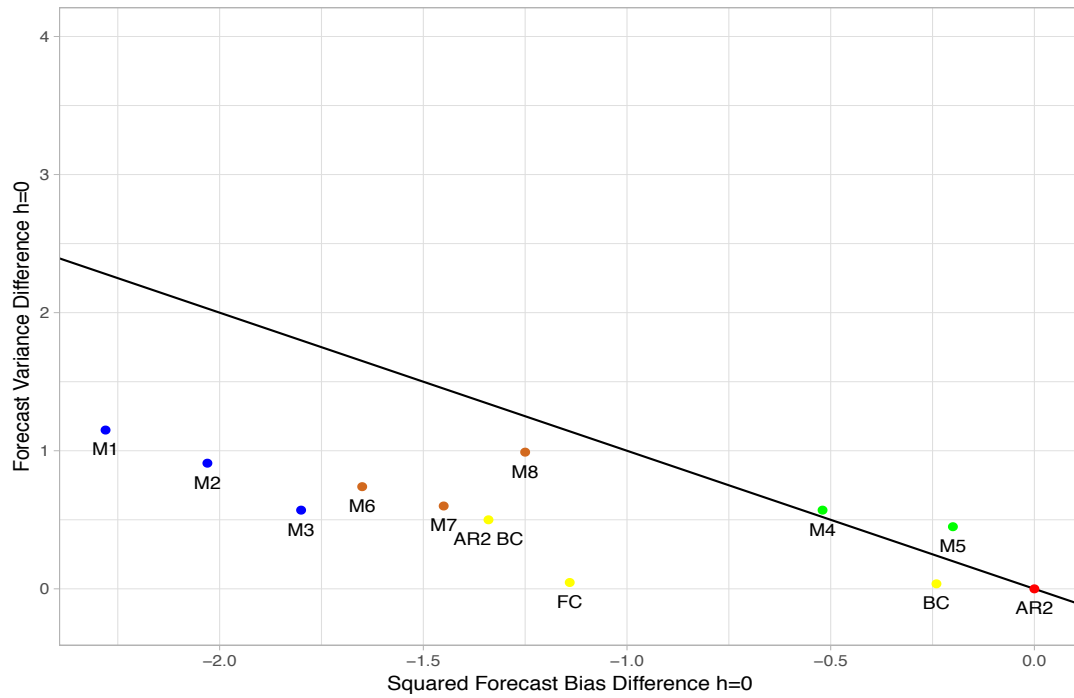


FIGURA 17: Gráfico de dispersão da variância e viés ao quadro da previsão relativa ao $AR(2)$, 1992T1-2017T2.

O eixo y e o eixo x representam a variância da previsão e o viés ao quadro para $h = 0$, calculados como a diferença entre a variância da previsão (viés ao quadrado) de $f_{t+h,t}^j$ a variância da previsão (viés ao quadrado) $AR(2)$. Cada ponto sobre a linha pontilhada representa uma previsão com o mesmo $MSFE$ que $AR(2)$; pontos a direita são previsões que são piores que o $AR(2)$, e pontos a esquerda são previsões melhores que o $AR(2)$. Em azul temos os modelos com o dicionários que variam no tempo; em verde temos os modelos com dicionário fixo; em amarelo temos os modelos de combinação de previsão e o Blue Chip; em laranja temos os modelos baseados em outros preditores; e em vermelho temos o benchmark Modelo $AR(2)$.

TABELA 15: Termos mais Selecionados.

Termos	Amostra Total(%)	Expansão(%)	Recessão(%)
recoveri	95.10	94.51	100.00
economic activity	60.78	56.04	100.00
weak	29.41	26.37	54.55
oil	25.49	27.47	9.09
upward pressure	18.63	16.48	36.36

Table 15 mostra a frequência relativa dos termos classificados como top-10 em termos de poder preditivo para todo o período fora da amostra e $h = 1$. Ranqueamos o poder preditivo das palavras pelo valor absoluto dos coeficientes sobre X_t , estimados por Elastic Net. Então os dez maiores coeficientes corresponda aos 10 maiores termos/palavras com maior poder preditivo.

TABELA 16: Previsões do Green Book do FED e dos preditores baseados nas minutas do FED

Modelo	h=0	h=1	h=3	h=6
Livro Verde do FED	0.732*	0.901**	1.041***	1.047
M1 Elastic Net	0.714*	0.759**	0.738*	0.732*

Tabela 16 mostra $RMSFE$ relativo ao benchmark $AR(2)$. O período fora da amostra é de 1992(T1) até 2012(T4), porque as previsões do Livro Verde são divulgadas com 5 anos de atraso. Para comparação reportamos o $RMSFE$ do M_1 para o mesmo período.

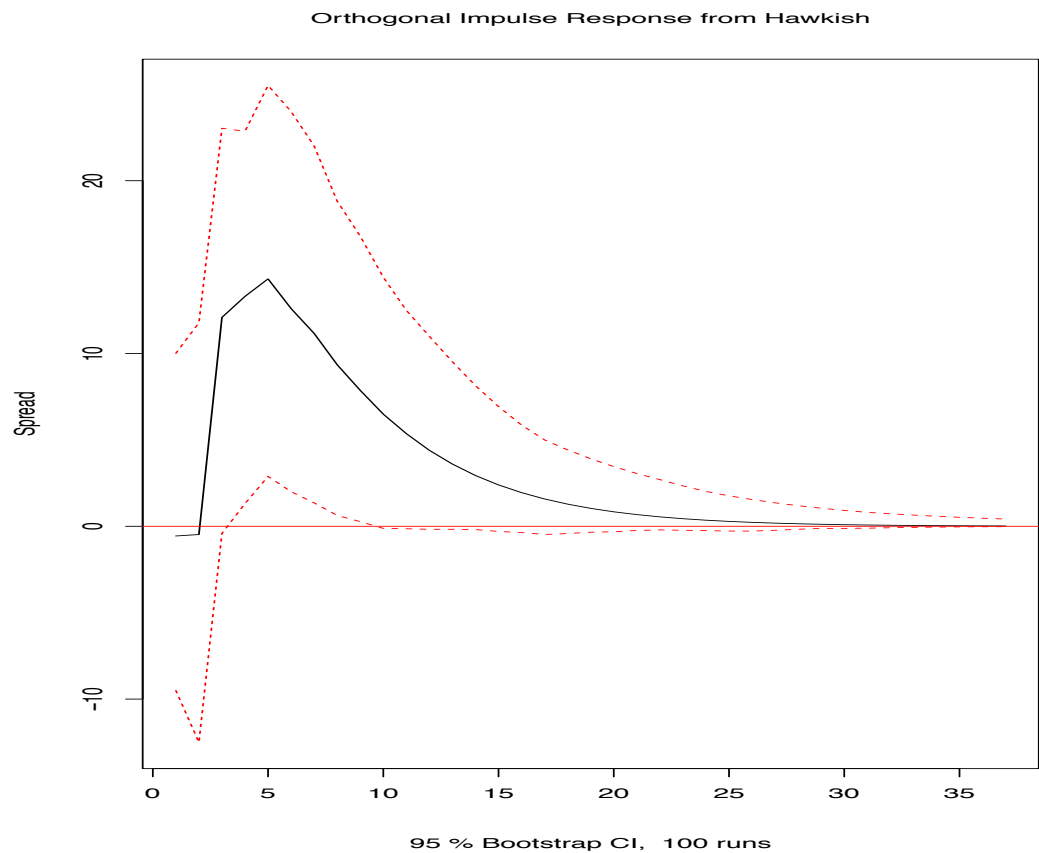


FIGURA 18: Resposta do Yield Spread a um choque no fator Hawkish.

Estimamos um modelo VAR ordenado como: {Hawkish Factor, Dovish Factor, Spread}. Nossa amostra inicia em 1984T1 e termina em 2017T2. Os fatores vem da metodologia 3. Seleccionamos os bigramas Hawkish e Dovish mais preditivos usando Elastic Net, então usamos PCA para computar o conjunto de fatores Hawkish e Dovish. Em seguida escolhemos o primeiro fator Hawkish e o primeiro fator Dovish. O yeild spread mede a diferença entre o rendimento de um título de 10 anos e 3 meses do tesouro americano(T10Y3M). Em () temos o símbolo da série de acordo com o FRED . Baseado no critério AIC, incluímos 2 defasagens no VAR.

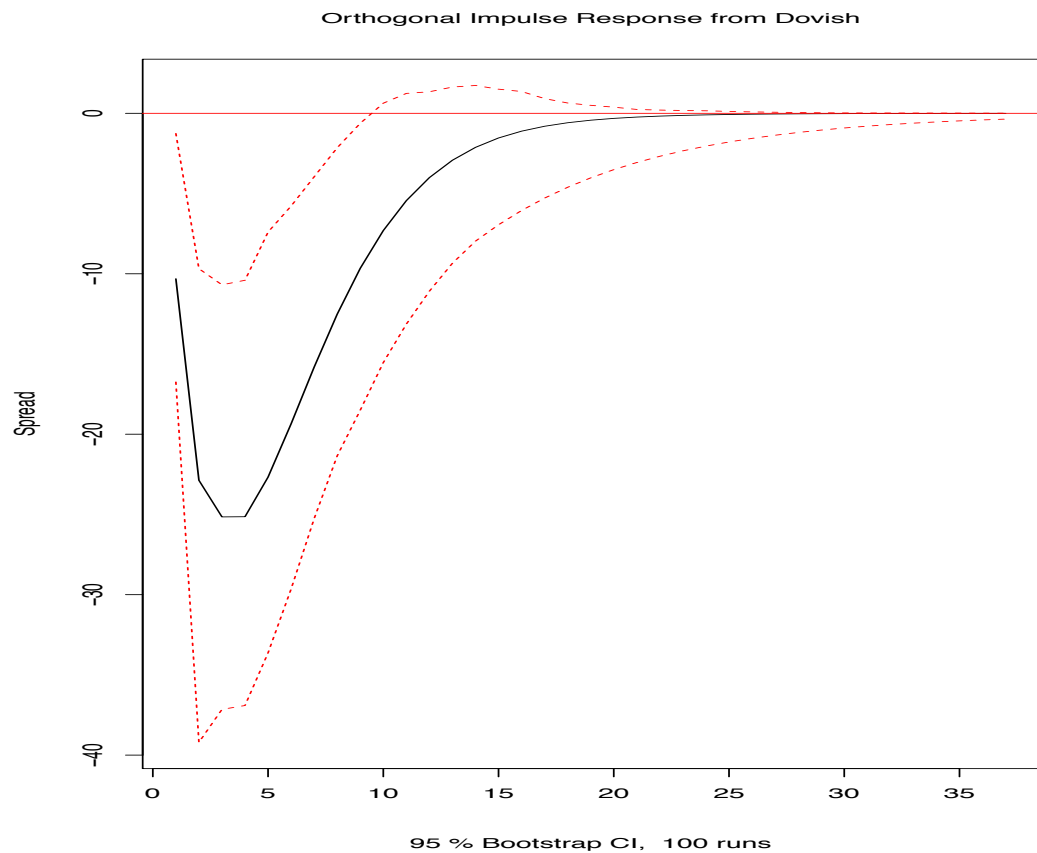


FIGURA 19: Resposta do the Yield Spread a um choque no fator Dovish.

Estimamos um modelo VAR ordenado como: {Hawkish Factor, Dovish Factor, Spread}. Nossa amostra inicia em 1984T1 e termina em 2017T2. Os fatores vem da metodologia 3. Seleccionamos os bigramas Hawkish e Dovish mais preditivos usando Elastic Net, então usamos PCA para computar o conjunto de fatores Hawkish e Dovish. Em seguida escolhemos o primeiro fator Hawkish e o primeiro fator Dovish. O yield spread mede a diferença entre o rendimento de um título de 10 anos e 3 meses do tesouro americano(T10Y3M). Em () temos o símbolo da série de acordo com o FRED . Baseado no critério AIC, incluímos 2 defasagens no VAR.

Apêndice C - Previsão do Prêmio de Risco: Identificando o Poder Preditivo das Notícias Financeiras.

TABELA 17: Resultado fora da amostra para a previsão do Prêmio de Risco.

Modelo	OOS:1999:1-2017:12				OOS:2008:1-2017:12			
	$R_2^{OS}(\%)$	DM	CW	$\Delta(anual\%)$	$R_2^{OS}(\%)$	DM	CW	$\Delta(anual\%)$
Modelos de Preditores Individuais								
DP	-0.60	0.95	0.69	0.46	-0.16	0.90	0.61	-0.34
DY	-0.53	0.95	0.64	0.62	-0.04	0.88	0.48	-0.05
EP	-0.86	0.93	0.58	1.89	-1.53	0.89	0.62	2.06
DE	-2.14	0.84	0.68	0.25	-3.06	0.82	0.63	0.72
SVAR	2.75	0.27	0.08	2.18	4.75	0.23	0.08	2.72
BM	-0.66	0.93	0.93	-0.41	-0.06	0.97	0.76	-0.08
NTIS	-1.62	0.85	0.96	-1.50	-2.21	0.97	0.97	-3.31
TBL	-1.87	0.38	0.83	-2.08	-0.28	0.10	0.43	-0.72
LTY	-1.73	0.41	0.85	-1.00	0.05	0.13	0.31	0.45
LTR	-0.29	0.73	0.44	-0.79	0.05	0.56	0.37	-1.24
TMS	-0.54	0.56	0.57	-0.17	-0.39	0.24	0.64	-1.23
DFY	-0.44	0.79	0.59	0.25	-0.73	0.74	0.59	0.32
DFR	-2.17	0.81	0.42	1.84	-1.72	0.77	0.37	0.59
INFL	-3.64	0.53	0.81	-1.63	-4.66	0.95	0.95	-5.60
E10P	-0.97	0.95	0.96	-0.67	-0.04	0.69	0.64	-0.01
Modelos Baseados em Notícias								
SVAR News	7.31	0.12	0.01	3.51	11.42	0.08	0.02	3.77
News	6.95	0.10	0.01	2.78	10.81	0.08	0.02	3.57
News ridge	4.05	0.39	0.05	1.37	6.59	0.28	0.07	0.43
News lasso	3.88	0.26	0.01	3.42	6.60	0.22	0.01	6.05
News factor	-0.95	0.92	0.51	0.23	0.63	0.63	0.25	0.70
Outros Modelos								
Sum-of-the-parts	1.09	0.02	0.04	2.85	0.56	0.10	0.11	2.00
Sum-of-the-parts News	4.90	0.02	0.00	7.62	6.27	0.05	0.02	9.98
FC Single Predictors	-0.26	0.62	0.60	0.48	0.17	0.54	0.40	1.48

Esta tabela reporta o R_2^{OS} (em %) e sua significância através dos p valores dos testes [Clark and West \(2007\)](#) (CW). Também reporta o p valor do teste [Diebold and Mariano \(2002\)](#) (DM) e o ganho de utilidade anual $\Delta(\%)$ associado a cada modelo para o período fora da amostra. O $R_2^{OS} > 0$, se a previsão condicional supera o benchmark fora da amostra. O ganho de utilidade anula é interpretado como a taxa de administração que um investidor estaria disposto a pagar para ter acesso a informação adicional do modelo condicional. FC denota a combinação de previsão dos preditores individuais.

TABELA 18: Resultado fora da amostra para a previsão do Prêmio de Risco para os sub períodos de expansão e recessão.

Modelo	OOS:Expansão				OOS:Recessão			
	$R_2^{OS}(\%)$	DM	CW	$\Delta(anual\%)$	$R_2^{OS}(\%)$	DM	CW	$\Delta(anual\%)$
Modelo de Preditores Individuais								
DP	-1.20	0.97	0.79	-0.05	0.57	0.16	0.16	5.19
DY	-1.27	0.97	0.80	-0.05	0.94	0.10	0.06	4.86
EP	-0.82	0.91	0.72	-0.05	-0.93	0.79	0.48	2.42
DE	-1.00	0.86	0.95	-0.05	-4.39	0.75	0.59	2.04
SVAR	1.11	0.18	0.14	6.86	5.96	0.37	0.12	-4.98
BM	-0.85	0.87	0.90	-0.05	-0.29	0.94	0.80	-6.82
NTIS	-0.72	0.38	0.76	-0.80	-3.40	0.96	0.94	-2.52
TBL	-0.89	0.17	0.46	-0.05	-3.82	0.93	0.97	-9.41
LTY	-1.20	0.20	0.59	-0.05	-2.76	0.96	0.97	-9.44
LTR	-0.55	0.64	0.52	0.82	0.21	0.72	0.39	-6.78
TMS	-0.50	0.60	0.48	11.53	-0.59	0.30	0.69	6.90
DFY	-0.24	0.64	0.89	0.27	-0.83	0.76	0.53	2.53
DFR	-1.97	0.69	0.45	12.66	-2.56	0.76	0.43	10.95
INFL	-1.02	0.23	0.28	10.21	-8.80	0.91	0.94	-1.33
E10P	-1.19	0.88	0.92	-0.05	-0.50	0.94	0.84	-8.92
Modelos Baseados em Notícias								
SVAR News	-0.44	0.66	0.37	-0.05	22.65	0.06	0.01	22.73
News	0.52	0.42	0.21	0.00	19.66	0.09	0.01	20.65
News ridge	-1.39	0.77	0.56	0.36	14.83	0.27	0.04	16.73
News lasso	2.05	0.34	0.05	0.00	7.49	0.30	0.03	10.80
News factor	-1.92	0.91	0.60	-0.07	0.99	0.64	0.25	-8.24
Outros Modelos								
Sum-of-the-parts	2.50	0.01	0.01	8.73	-1.71	0.75	0.94	9.61
Sum-of-the-parts News	3.30	0.01	0.00	8.83	8.05	0.23	0.04	22.02
FC Single Predictors	-0.25	0.33	0.68	-0.05	-0.26	0.79	0.52	-5.39

Esta tabela reporta o R_2^{OS} (em %) e sua significância através dos p valores dos testes [Clark and West \(2007\)](#) (CW). Também reporta o p valor do teste [Diebold and Mariano \(2002\)](#) (DM) e o ganho de utilidade anual $\Delta(\%)$ associado a cada modelo para o período fora da amostra. O $R_2^{OS} > 0$, se a previsão condicional supera o benchmark fora da amostra. O ganho de utilidade anula é interpretado como a taxa de administração que um investidor estaria disposto a pagar para ter acesso a informação adicional do modelo condicional. FC denota a combinação de previsão dos preditores individuais. As datas de recessão de acordo com a NBER são 2001:3-2001:11 e 2007:12-2009:6.

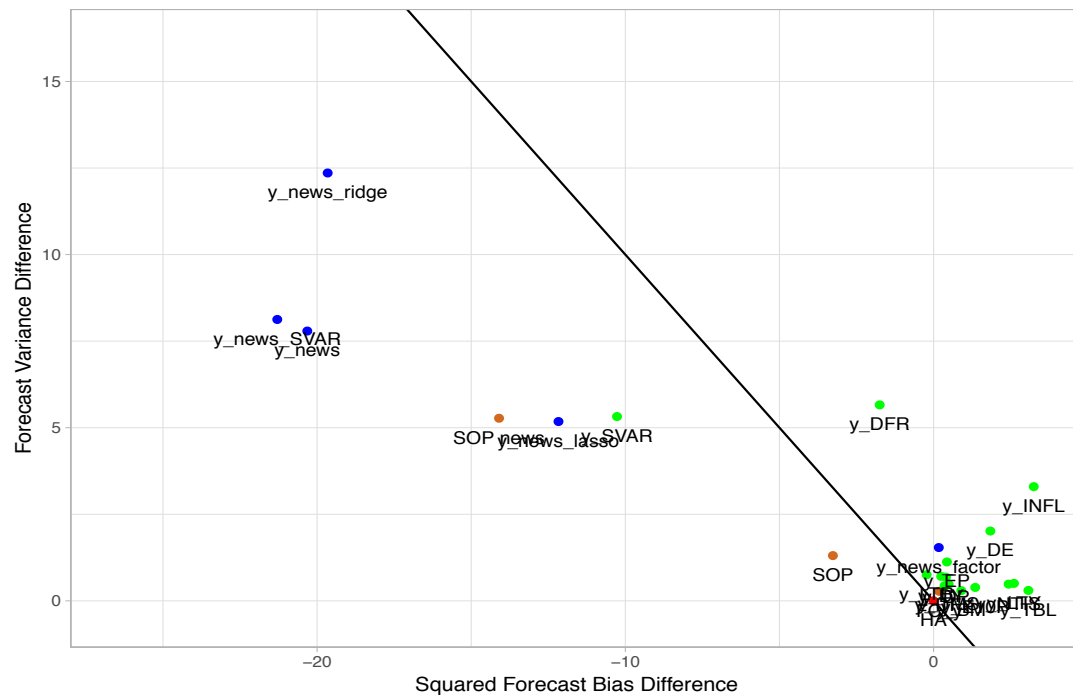


FIGURA 20: Gráfico de dispersão da variância da previsão e do viés ao quadrado relativo a média histórica, 1999:1–2017:12.

O eixo y e o eixo x representam a variância e o viés ao quadrado relativo à média histórica dos modelos de previsão, calculados como a diferença entre a variância da previsão (viés ao quadrado) dos modelos condicionais e da média histórica HA. Cada ponto sobre a linha representa uma variância igual a HA.; pontos à direita são previsões piores que HA ao passo que pontos à esquerda são previsões melhores que HA. O ponto azul representa o grupo news Models, o verde preditores individuais, o laranja outros preditores e o vermelho HA.

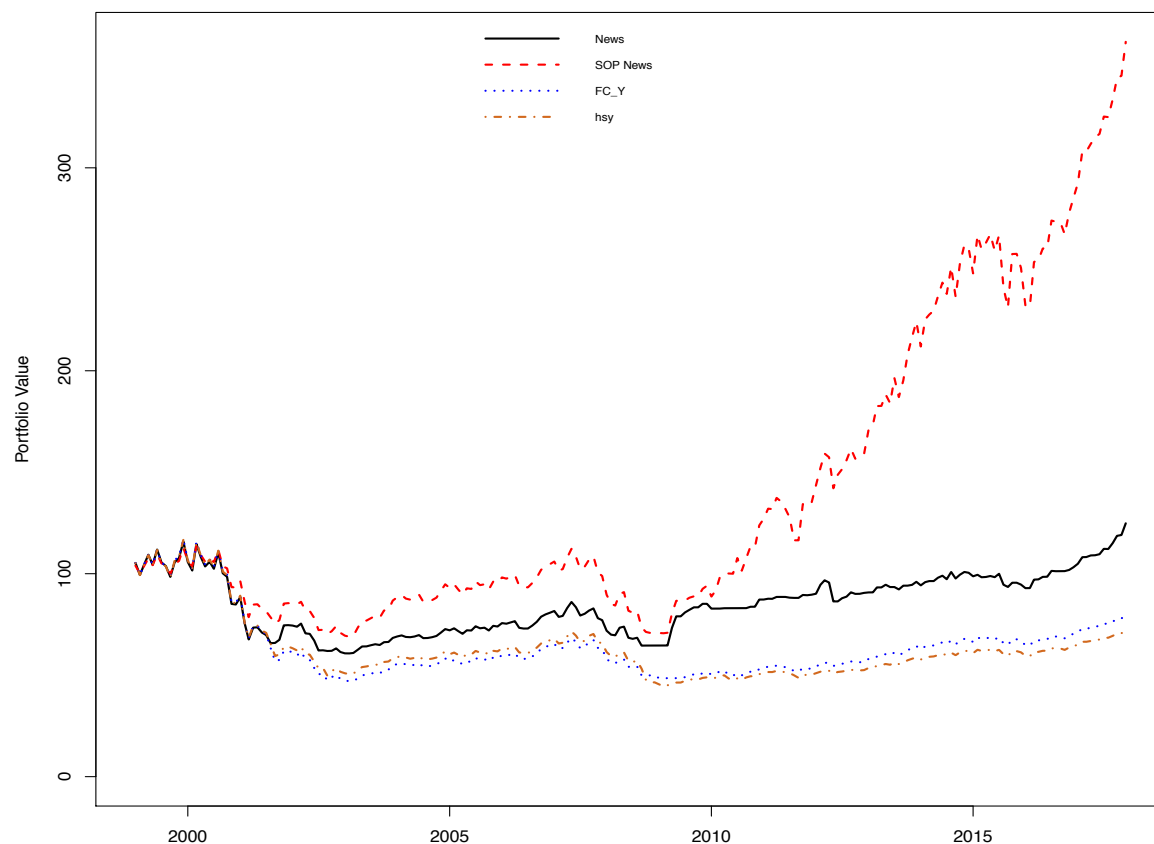


FIGURA 21: Riqueza acumulada associada a cada modelo de previsão 1999:1–2017:12.

A figura representa o valor dos portfólios selecionados com base no modelos de previsão ao longo do tempo. A figura mostra como um portfólio que valia U\$\$ 100.00 evoluiu no tempo e quanto ele vale em 2017:12.

TABELA 19: Palavras/termos mais preditivos.

Termo	OOS:1999:1-2017:12(%)	OOS:2008:1-2017:12(%)
world economy	100.00	100.00
banking system	89.47	96.64
junk bond	71.05	80.67
crude oil	64.91	78.15
retail sale	55.26	95.80
credit market	48.25	92.44
imf	37.72	14.29
economic growth	13.16	0.00
stock option	13.60	0.00
technology stock	2.19	0.84

Tabela 19 mostra a frequência relativa das raízes de palavras/termos classificados como top-10 in termos de poder preditivo para todo período fora da amostra. Ranqueamos o poder preditivo das palavras usando o valor absoluto dos coeficientes estimados por Elastic Net. Então os 10 maiores coeficientes correspondem as palavras mais preditivas.

Índice Remissivo

A

- Análise de Componente Principal, [46](#)
- Aprendizado de Máquina, [54](#)
- Aprendizado de Máquina não Supervisionado, [33](#)
- Aprendizado de Máquina Supervisionado, [13](#), [32–35](#), [50](#), [72](#)

E

- Elastic Net, [32](#), [35](#), [38](#), [40](#), [41](#), [43–48](#), [51–54](#), [58](#), [64–66](#), [68](#), [72](#)

I

- Incerteza Macroeconômica, [15–17](#), [19](#), [22](#), [23](#), [25](#), [26](#), [29](#)

L

- LASSO, [35](#), [41](#), [45](#), [46](#), [52](#), [60](#), [66](#)

N

- News Models, [69](#)
- Notícias Financeiras, [14](#), [62–66](#), [68–72](#), [74](#)

R

- Regressão de Texto, [35](#), [50](#)
- Ridge Regression, [35](#), [45](#), [46](#), [52](#)